

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

AE

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: <b>C12N 15/12, C07K 14/47, G01N 33/68, C12N 15/62, C07K 16/18</b>		A2	(11) International Publication Number: <b>WO 99/06545</b> (43) International Publication Date: <b>11 February 1999 (11.02.99)</b>
(21) International Application Number: <b>PCT/EP98/04811</b> (22) International Filing Date: <b>31 July 1998 (31.07.98)</b>		(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: <b>97 11 3306.1 1 August 1997 (01.08.97) EP</b>		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(71) Applicant (for all designated States except US): <b>MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. (DE/DE); Berlin (DE).</b>			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): <b>WANKER, Erich [AU/DE]; Leichhardtstrasse 61, D-14195 Berlin (DE). LEHRACH, Hans [AU/DE]; Lützelsteiner Weg 50, D-14195 Berlin (DE). SCHERZINGER, Eberhard [DE/DE]; Lützelsteiner Weg 52, D-14195 Berlin (DE). BATES, Gillian [GB/GB]; Division of Medical and Molecular Genetics, United Medical and Dental Schools, Guy's Hospital, Guy's Tower, 8th floor, London SE1 9RT (GB).</b>			
(74) Agent: <b>VOSSIUS &amp; PARTNER; Postfach 86 07 67, D-81634 München (DE).</b>			
(54) Title: <b>COMPOSITION AND METHOD FOR THE DETECTION OF DISEASES ASSOCIATED WITH AMYLOID-LIKE FIBRIL OR PROTEIN AGGREGATE FORMATION</b>			
<p><b>GST-HD20AP:</b> GST-IBGRGIR-MATLEKLMKAFESLKSFIQ<sub>n</sub>IPPPP-LERPHRD</p> <p><b>GST-HD51AP:</b> GST-IBGRGIR-MATLEKLMKAFESLKSFIQ<sub>n</sub>IPPPP-LERPHRD</p> <p><b>GST-HD20APBio:</b> GST-IBGRGIR-MATLEKLMKAFESLKSFIQ<sub>n</sub>IPPPP-LERPHRD</p> <p><b>GST-HD51APBio:</b> GST-IBGRGIR-MATLEKLMKAFESLKSFIQ<sub>n</sub>IPPPP-LERPHRD</p>			
(57) Abstract			
<p>The present invention relates to novel compositions useful for elucidating the onset or progress of diseases of preferably neuronal origin associated with the formation of amyloid-like fibrils or protein aggregates. Further, the present invention relates to methods for monitoring said formation as well as to methods for identifying inhibitors of said formation. Additionally, the invention relates to inhibitors identified by the method of the invention as well as to pharmaceutical compositions comprising said inhibitors.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KR	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**composition and method for the detection of diseases associated  
with amyloid-like fibril or protein aggregate formation**

The present invention relates to novel compositions useful for elucidating the onset or progress of diseases of preferably neuronal origin associated with the formation of amyloid-like fibrils or protein aggregates. Further, the present invention relates to methods for monitoring said formation as well as to methods for identifying inhibitors of said formation. Additionally, the invention relates to inhibitors identified by the method of the invention as well as to pharmaceutical compositions comprising said inhibitors.

A variety of diseases, both in humans and animals, is characterized by the pathogenic formation of amyloid-like fibrils or protein aggregates in neuronal tissues. A well-known and typical example of such diseases is Alzheimer's disease (AD). AD is characterized by the formation of neurofibrillar tangles and  $\beta$ -amyloid fibrils in the brain of AD patients. Similarly, scrapie is associated with the occurrence of scrapie-associated fibrils in brain tissue.

Another class of these diseases is characterized by an expansion of CAG repeats in certain genes. The affected proteins display a corresponding polyglutamine expansion. Said diseases are further characterized by a late onset in life and a dominant pathway of inheritance.

A typical representative of this class of diseases is Huntington's disease. Huntington's disease (HD) is an autosomal dominant progressive neurodegenerative disorder characterized by personality changes, motor impairment and subcortical dementia (Harper, 1991). It is associated with a selective neuronal cell death occurring primarily in the cortex and striatum (Vonsattel et al., 1985). The disorder is caused by a CAG/polyglutamine (polyGln) repeat expansion in the first exon of a gene encoding a large ~350 kDa protein of unknown function and designated

huntingtin (HDCRG, 1993). The CAG repeat is highly polymorphic and varies from 6-39 repeats on chromosomes of unaffected individuals and 35-180 repeats on HD chromosomes (Rubinsztein et al., 1996; Sathasivam et al., 1997). The majority of adult onset cases have expansions ranging from 40-55 units, whereas expansions of 70 and above invariably cause the juvenile form of the disease. The normal and mutant forms of huntingtin have been shown to be expressed at similar levels in the central nervous system and in peripheral tissues (Trottier et al., 1995a). Within the brain, huntingtin was found predominantly in neurons and was present in cell bodies, dendrites and also in the nerve terminals. Immunohistochemistry, electron microscopy and subcellular fractionations have shown that huntingtin is primarily a cytosolic protein associated with vesicles and/or microtubules, suggesting that it plays a functional role in cytoskeletal anchoring or transport of vesicles (DiFiglia et al., 1995; Gutekunst et al., 1995; Sharp et al., 1995) Huntingtin has also been detected in the nucleus (de Rooij et al., 1996; Hoogeveen et al., 1993) suggesting that transcriptional regulation cannot be ruled out as a possible function of this protein.

In addition to HD, CAG/polyglutamine expansions have been found in at least seven other inherited neurodegenerative disorders which include: spinal and bulbar muscular atrophy (SBMA), dentatorubral pallidoluysian atrophy (DRPLA), and the spinocerebellar ataxias (SCA) types 1, 2, 3, 6 and 7 (referenced in Bates et al. 1997). The normal and expanded size ranges are comparable with the exception of SCA6 in which the expanded alleles are smaller and the mutation is likely to act by a different route. However, in all cases the CAG repeat is located within the coding region and is translated into a stretch of polyglutamine residues. Although the proteins harbouring the polyglutamine sequences are unrelated and mostly of unknown function, it is likely that the mutations act through a similar mechanism. Without exception, these proteins are widely expressed and generally localized in the cytoplasm. However, despite overlapping expression patterns in brain, the neuronal cell death is relatively specific and can differ markedly (Ross, 1995), indicating that additional factors are needed to convey the specific patterns of neurodegeneration.

Several investigators have proposed that HD is caused by a toxic gain of function, which in turn is caused by abnormal protein-protein interactions related to the elongated polyglutamine. It is possible that the binding of a protein to the polyglutamine region could either confer a new property on huntingtin or alter its normal interactions causing selective cell death either through the specific expression patterns of the interacting protein or through the selective vulnerability of certain cells. To date, four

potential huntingtin-interacting proteins have been isolated: HAP1 (Li et al., 1995), GAPDH (Burke et al., 1996), HIP2 (Kalchman et al., 1996) and HIP-1 (Kalchman et al., 1997; Wanker et al., 1997). However, it has not been demonstrated whether the binding of these proteins to huntingtin is involved in the selective neuropathology. A gain of function mechanism has been supported by the identification of an antibody that specifically reacts with the pathogenic polyglutamine expansions (Trottier et al., 1995b). This indicates that upon expansion into the pathogenic range, a polyglutamine sequence may undergo a conformational change. Poly-L-glutamines form pleated sheets of  $\beta$ -strands held together by hydrogen bonds between their amides (Perutz et al., 1994). It was proposed that the expanded glutamine repeats in huntingtin may function as polar zippers, joining protein molecules together (Perutz, 1996). In the long run, this could result in the precipitation of huntingtin protein in specific neurons causing the observed selective neuronal loss. Thus, the mechanism underlying HD would be similar to scrapie, Creutzfeldt-Jakob or Alzheimer's disease, in which  $\beta$ -sheet secondary structures lead to the formation of toxic protein aggregates in selective neurons (Caughey and Chesebro, 1997).

Recently, strains of mice (R6) that are transgenic for the HD mutation have been generated (Mangiarini et al., 1996). In these mice exon 1 of the human HD gene carrying CAG repeat expansions of 115-156 units is expressed under the control of the human HD promoter. It has been demonstrated that the transgenic animals exhibit a progressive neurological phenotype that exhibits many of the motor and non motor features of HD. The phenotype includes a resting tremor; irregular gait; rapid, abrupt shuddering movements; stereotypic grooming movements and epileptic seizures. Coincident with the onset of the movement disorder the mice show a progressive weight loss. Neuropathological analysis has shown a reduction in brain weight (which precedes that in body weight) and the presence of neuronal intranuclear inclusions (NIIs) predating any evidence of neuronal dysfunction (Davies et al., 1997). The NIIs are immunoreactive for N-terminal huntingtin antibodies that detect the transgene protein and for ubiquitin but do not contain the endogenous mouse huntingtin. At the ultrastructural level, a solitary intranuclear inclusion appears as a roughly circular pale structure of a fine granular nature with occasional filamentous structures and devoid of a membrane. In addition, the neurons invariably have indentations of the nuclear membrane and an apparent increase in the density and clustering of nuclear pores. All three of these ultrastructural nuclear changes have previously been reported in EM studies from HD patients (Roizman et al., 1979; Roos and Bots, 1983; Tellez-Nagel et al., 1974).

Thus, a large body of data has accumulated that describes aspects of the pathology of the above-discussed diseases. However, the actual mechanisms leading to the onset of the various disease states are still unknown. Although a variety of hypotheses have been formulated in the art, it is equally unknown how the amyloid or aggregate formation is triggered or caused within affected cells or tissues. Without a detailed knowledge of the formation of said aggregates, the development of a suitable pharmaceutical composition for treating such diseases appears rather difficult. The technical problem underlying the present invention was therefore to provide means and methods suitable for the eventual elucidation of the etiology of these diseases and the development of appropriate medicines.

The above technical problem is solved by the embodiment characterized in the claims:

Accordingly, the present invention relates to a composition comprising

- (a) a nucleic acid molecule encoding a fusion protein comprising
- (aa) a (poly)peptide that enhances solubility and/or prevents aggregation of said fusion protein; and
- (ab) an amyloidogenic (poly)peptide that has the ability to self-assemble into amyloid-like fibrils or protein aggregates;
- (b) a vector containing the nucleic acid molecule of (a);
- (c) a host transformed with the vector of (b);
- (d) a fusion protein encoded by the nucleic acid of (a) or a functional derivative thereof; and/or
- (e) an antibody specific for the fusion protein of (d).

As used herein, the term "(poly)peptide" relates to a polypeptide or a peptide depending on the length of the amino acid string. Said (poly)peptide has the ability to enhance solubility of a fusion partner in said fusion protein and thus of the fusion protein itself. Additionally, or alternatively, said (poly)peptide prevents the aggregation of the fusion partner and thus of the fusion protein. Said (poly)peptide is combined within said fusion protein with an amyloidogenic (poly)peptide having the above recited features. The connection of both (poly)peptides may be via a linker or by a direct attachment. It is preferred that either the linker or either (poly)peptide comprises a cleavable site. Said cleavable site should render both (poly)peptides essentially intact. Alternatively, said fusion protein may comprise a number of cleavage sites. Upon cleavage, which may be exhaustive or under limiting conditions, the amyloidogenic (poly)peptide should, when used for the purposes of

the present invention, retain the ability to self-assemble. The person skilled in the art is in the position to determine appropriate conditions for a corresponding limited cleavage. The composition of the invention may comprise one, several, or all of the compounds recited in features (a) to (e), above.

The term "functional derivative" refers to a fusion protein which comprises, for example, modified amino acids or amino acid substitutions and retains the functions of the fusion protein detailed herein above.

The term "antibody specific for the fusion protein" comprised in the composition of the invention is intended to mean that said antibody is only specific for the fusion protein but not for either of the above cited components of said fusion protein.

In accordance with the present invention, it could surprisingly be shown that the composition comprising the above recited components can be used for the elucidation of amyloid-like fibril or protein aggregate formation. The components of the composition can be used in varying combinations to test, for example, for specific conditions under which amyloids are formed *in vitro*. The *in vitro* data obtained with the composition of the invention may then be compared to or brought into relation with the *in vivo* situation and appropriate conclusions may be drawn therefrom.

The *in vitro* systems that can be established with the composition of the invention allow formation of highly stable amyloid-like protein aggregates. Such aggregates may be obtained, for example, by proteolytic cleavage of GST fusion proteins comprising exon 1 of the HD gene and containing expanded polyglutamine sequences. Alternatively, such aggregates may be obtained by lowering the pH value from 8 to 5 or by increasing the protein concentration. The arrays of fibrillar structures of varying sizes and shapes observed by electron microscopy surprisingly clearly resemble those of purified amyloids. Furthermore, the polarization microscopic properties of the fibrils stained with Congo red are strikingly similar to those described for amyloids. The green-gold birefringence of the amyloid-like fibrils indicates that the polymers have common structural features. Although the Congo red staining does not determine conclusively whether the fibrils consist of  $\beta$ -pleated sheets, the method suggests that this is likely in view of experience gained with other protein polymers (Caputo et al., 1992). However, it has been generally accepted that naturally occurring mammalian protein polymers that exhibit fibrillar structures and green birefringence after Congo red staining should be classified as amyloids (Glenner, 1980). Instead of the HD gene, other nucleic acid sequences encoding

amyloidogenic (poly)peptides may be used to generate said fusion proteins. Preferably, the composition of the invention is a diagnostic composition. The composition of the invention may also be a kit. The diagnostic composition can advantageously be employed in the assessment of a disease state whereas the kit may rather be employed in the development of, for example, inhibitors or in the elucidation of amyloid formation.

It is a preferred embodiment of the composition of the invention, that said amyloidogenic (poly)peptide comprises a polyglutamine expansion. In the prior art it has been shown by X-ray diffraction studies that synthetic peptides containing polyglins form  $\beta$ -sheets strongly held together by hydrogen bonds (Perutz et al., 1994). Because synthetic poly(L-glutamine) is insoluble in water, a synthetic peptide with the sequence Asp<sub>2</sub>-Gln<sub>15</sub>-Lys<sub>2</sub> was used in that study. A stretch of 10 glutamines was also inserted into the loop of chymotrypsin inhibitor-2 (CI2), and it was demonstrated by analytical ultracentrifugation that the recombinant protein, in addition to monomers formed dimers and trimers, whereas wild-type CI2 was present only in the monomeric form (Stott et al., 1995). It has been proposed that the polyglin stretch functions as a polar zipper, joining proteins together. However, the hypothesis that glutamine repeats in proteins form  $\beta$ -pleated sheets and induce protein aggregation by a mechanism similar to that observed in spongiform encephalopathy (TSE) diseases (Caughey and Chesebro, 1997) could not be proven with this recombinant protein. Most likely, the length of the polyglin sequence inserted into CI2 was too short. Accordingly, the experimental data actually obtained teach away from the above recited hypothesis.

In a particularly preferred embodiment said polyglutamine expansion comprises at least 35 glutamines. In a further particularly preferred embodiment said polyglutamine expansion comprises at least 51 glutamines.

Our studies with the GST-HD fusion proteins containing polyglin sequences of varying lengths demonstrate that a certain length of the polyglin stretch is necessary for the formation of amyloid-like fibrils *in vitro*. When the purified fusion proteins were analyzed by SDS-PAGE, insoluble high molecular weight protein aggregates were only detected with the proteins containing 81 and 122 glutamines (Fig. 1a and b), whereas the protein with 51 glutamines was soluble and no fibrillar structures were detected by electron microscopy (Fig. 4a). This indicates that the critical length of the polyglin stretch in the fusion proteins leading to the formation of aggregates is greater than 51 glutamines. Accordingly, fusion proteins of the invention with a

polyglutamine expansion of more than 51 glutamines, such as 81 or 122 glutamines, may be employed in studies for the formation of aggregates that render a cleavage reaction unnecessary. However, when the GST-tag, which is known to enhance the solubility of many proteins (Smith and Johnson, 1988), was cleaved by limited digestion with trypsin, the liberated HD exon 1 protein with 51 glutamines also started to form aggregates (Fig. 3a) and the amount of these aggregates increased when the GST-tag was totally degraded with trypsin (Fig. 3b). This indicates that in the HD exon 1 protein 51 glutamines are sufficient to form aggregates, whereas 20 and 30 glutamines under the same conditions are not. The minimum critical length essential for the development of amyloid-like structures after removal of the GST-tag is not known and has to be determined. However, preliminary experiments in our laboratory suggest that the threshold for the formation of HD exon 1 protein aggregates is between 35 - 48 glutamines. This result is strikingly similar to the pathological threshold in HD, SBMA, DRPLA, SCA1, SCA2, SCA3 and SCA7. In all of these neurodegenerative polyglutamine diseases a pathological phenotype was found when more than 41 repeats were present, suggesting that the elongation of the polyglutamine repeat beyond a certain length may lead to a phase change in the affected proteins. This could, for example, be a change from random coils to hydrogen-bonded hairpins in the polyglutamine stretch, see Perutz (1996).

With the understanding that the applicant is not bound by any scientific theory, a mechanism is proposed for the fibril formation induced by proteolytic cleavage of GST-HD51 as shown in Fig. 7. Based on the known crystal structure of GST with a C-terminal fusion peptide (Lim et al., 1994) and the fact that the purified GST protein is a dimer we suppose that native GST-HD51 exists as a dimer with two expanded polyglutamine sequences which form stable hairpins consisting of antiparallel  $\beta$ -strands strongly held together by hydrogen bonds between the main chain and the side chain amides. In the native protein both hairpins are tightly bound to the surface of GST and not accessible for protein-protein interactions with other polyglutamine sequences. As a result of the cleavage with a site-specific protease, both hairpins become accessible and  $\beta$ -sheets with hairpins from other cleaved protein molecules are formed. This transient population of intermediates consisting of GST molecules and hairpins leads to the formation of polyglutamine-containing  $\beta$ -sheet fibrils and free GST molecules. This model is supported by the finding of potential intermediate structures present on one or both ends of the growing fibrils (Fig. 4c and d). These clots of varying sizes were not detected when GST-HD51 was digested to completion with trypsin, which totally degrades the GST-tag, whilst they were detectable upon limited

digestion, leaving the GST moiety largely intact (Fig. 4d). This indicates that these structures are transient intermediates.

A model of the formation of amyloid-like fibrils via transient intermediates is not without precedent. Booth et al. (1997) have shown that amyloidogenic lysosome variants aggregate on heating, unlike the wild-type protein, and that the lysozyme fibrils are formed from potential precursor proteins. It is possible that the transient GST-HD intermediates function as nuclei for ordered protein aggregation, very similarly to protein crystallization and microtubule formation, which are nucleation-dependent polymerisations (Jarrett and Lansbury, 1993). Once a nucleus is formed, the further addition of monomers becomes thermodynamically favorable and results in rapid polymerization. An important feature of a nucleation-dependent process is a lag time before the aggregates are detectable. During this period, dimers and trimers are formed. Fig. 3a shows that during proteolytic cleavage of GST-HD51 dimers of the released HD portion are formed, the concentration of which then decreases upon prolonged incubation concomitant with an increase in the formation of large protein aggregates (Fig. 3b). Although additional kinetic studies will be necessary to prove this assumption, preliminary results in our laboratory suggest that a "one-dimensional" crystallization leads to the formation of *in vitro* amyloid-like huntingtin aggregates.

Accordingly, the present invention provides both the possibilities to analyze aggregation of amyloid-like aggregates using defined cleavage conditions or using fusion proteins comprising long polyglutamine stretches that render the cleavage unnecessary. Whereas the cleavage of the fusion protein enables to set a distinct starting point of the reaction and therefore of the aggregate formation, the second alternative has the advantage that the use of a cleaving agent is rendered obsolete.

In a further particularly preferred embodiment of the invention said (poly)peptide defined in (ab) is huntingtin, androgen receptor, atropin, TATA binding protein, or ataxin-1,-2,-3,-6 or -7 or a fragment or derivative thereof.

The fibrillar structures formed by proteolytic cleavage of purified GST-HD51 *in vitro* and also in the brains of mice transgenic for the HD mutation are very similar to structures detected in brain sections or purified protein fractions of Alzheimer's disease (AD), Creutzfeldt-Jakob disease (CJD), Parkinson's Disease, Gerstmann-Strässler-Scheinker syndrome (GSS), fatal familial insomnia (FFI), kuru, bovine spongiform encephalopathy (BSE) and scrapie (Caughey and Chesebro, 1997). In all

these disorders, the accumulation of amyloid-like fibrils in the central nervous system is accompanied by loss of nerve cells and a neuropathological phenotype. However, the molecular basis of these diseases is not known. For the first time, our results raise the possibility that HD, DRPLA, SBMA, SCA1, SCA2, SCA3 and SCA7 are also the result of a toxic amyloid fibrillogenesis. Although the detection of amyloid-like fibrils has not previously been reported in these inherited diseases, our results strongly suggest that polyglu-containing polymers are also formed *in vivo* by their detection in a transgenic model of polyglutamine disease. The high molecular weight aggregates were exclusively detected in the nuclear fraction prepared from transgene brain material, which is in good agreement with the results of Davies et al. (1997), who demonstrated the presence of the transgene protein and ubiquitin in neuronal intranuclear inclusions, from a time prior to the development of a neurological phenotype. Strikingly, ultrastructural analysis has shown similar intranuclear inclusions to be present in the cortical and striatal biopsy material from HD patients (Roizin et al., 1979) some of which showed clear evidence of intranuclear fibrils of up to 1  $\mu$ m in length. Preliminary experiments with nuclear protein fractions prepared from HD brain material indicated that insoluble huntingtin aggregates are indeed present in these fractions. However, additional control experiments have to be performed to substantiate these results.

One possible explanation for the absence of detection of high molecular weight huntingtin protein aggregates in HD brains could be that the aggregates consist mainly of polyglu-containing peptides which have been cleaved from the full length protein. In such a case, only an antibody raised against an N-terminal huntingtin fragment, containing the polyglu sequence, would be able to detect the aggregates in the nucleus. In most of the previous immunohistochemical studies, antibodies raised against the central or C-terminal portion of huntingtin have been used, which detect the full length protein (350 kDa) in the cytosol and in the membrane containing fractions (DiFiglia et al., 1995; Sharp et al., 1995; Trottier et al., 1995a). However, antibodies raised against peptides and fusion proteins from the N- and C-terminus of huntingtin also detected the protein in the nucleus (de Rooij et al., 1996; Hoogeveen et al., 1993), indicating that huntingtin is also present in this subcellular compartment. There are several lines of evidence to implicate a shorter polyglu-containing peptide/protein fragment of huntingtin in the pathology of HD. Ikeda et al. (1996) showed that a short fragment of the MJD1 protein containing 79 polyglus (Q79C) but not the full length protein with the elongated repeat induced apoptotic cell death in COS cells. The polyglu-containing protein fragment migrated in SDS-gels at a position much higher than expected from its molecular weight, even after boiling in

the presence of 2% SDS. These results are in good agreement with our data obtained using the GST-HD fusion proteins containing elongated polyglutamine sequences. Fig. 1a and b show that the expression of GST-HD83 and GST-HD122 in *E. coli* was dramatically reduced compared to the fusion proteins containing 20-51 repeats, and additional studies have indicated that the elongated glutamines are toxic for *E. coli* cells.

The possibility that polyglutamine-containing cleavage products of huntingtin cause neurodegeneration in HD is substantiated by the finding of Goldberg et al. (1996) who showed that an N-terminal 80 kDa huntingtin fragment is cleaved from the full length protein by apopain, a proapoptotic cysteine protease. This indicates that the N-terminus of huntingtin is primarily accessible for proteases and distinct proteolytic cleavage products can be formed *in vitro* and *in vivo*. In addition, there is strong evidence that the mutated huntingtin somehow induces apoptotic cell death in HD, but the underlying molecular mechanism is not known (Duyao et al., 1995; Portera-Cailliau et al., 1995). Our data suggest that a proteolytic cleavage product of huntingtin, which is transported into the nucleus by an unknown mechanism, causes selective neuronal cell death by the formation of insoluble amyloid-like fibrils. It is possible that the transport to the nucleus is facilitated by a specific nuclear transport mechanism which is unique to certain neuronal cells and involves abnormal protein-protein interactions related to the elongated polyglutamine. Alternatively, there may be specific nuclear proteins in the affected neurons which enhance the huntingtin protein aggregation.

Recently, the formation of neuronal intranuclear inclusions in mice transgenic for the SCA1 mutation have been detected, indicating that polyglutamine-containing polymers are also formed in spinocerebellar ataxia type 1. Furthermore, the accumulation of polyglutamine-containing protein aggregates in neuronal intranuclear inclusions (NIIs) has been demonstrated for several progressive neurodegenerative diseases such as Huntington's disease (HD) (M. DiFiglia et al., 1997; M. W. Becher et al., 1997), dentatorubral pallidoluysian atrophy (DRPLA) (M. W. Becher et al., 1997; S. Igarashi et al., 1998) and the spinocerebellar ataxia (SCA) types 1 (P. J. Skinner et al., 1997; A. Matilla et al., 1997), 3 (H. L. Paulson et al., 1997) and 7 (M. Holmberg et al., 1998).

The components of the composition of the invention may be packaged in containers such as vials, optionally in buffers and/or solutions. If appropriate, one or more of said components may be packaged in one and the same container.

In an additional preferred embodiment of the composition of the present invention, said amyloidogenic (poly)peptide self-assembles subsequent to release from said fusion protein.

As has been pointed out herein before, the self-assembly of said (poly)peptides only subsequent to the release from the fusion protein provides the advantage that an exact time point of the start of the formation can be set. This has a number of advantages. For example, inhibitors of aggregate formation can be tested as regards their efficacy as a function of time. In addition, obtainment of data is facilitated in view of the fact that for amyloid formation a premix may be set up to which only the cleaving agent must be added.

In a further preferred embodiment of the composition of the invention, said amyloidogenic (poly)peptide is the amyloid precursor protein (APP),  $\beta$ -protein, an immunoglobulin light chain, serum amyloid A, transthyretin, cystatin C,  $\beta$ 2-microglobulin, apolipoprotein A-1, gelsoline, islet amyloid polypeptide (IAPP), calcitonin, a prion, atrial natriuretic factor (ANF), lysozyme, insulin, fibrinogen, tau proteins or  $\alpha$ -synuclein or a fragment or derivative thereof.

Deposits of  $\beta$ -amyloid in neuritic plaques and blood vessel walls are the principal pathological feature in the brains of patients with Alzheimer's disease. These amyloid deposits contain the 39-43 amino acid  $\beta$ -amyloid peptide which is derived by proteolytic cleavage from the larger precursor, the amyloid precursor protein (APP). There is strong evidence that the formation and aggregation of  $\beta$ -amyloids into fibrils is the primary pathogenic event leading to amyloid deposition in Alzheimer's disease.

An additional preferred embodiment relates to a composition wherein said (poly)peptide defined in (aa) is glutathione S-transferase (GST), intein, thioredoxin, dihydrofolate reductase (DHFR) or chymotrypsin inhibitor 2 (CI2) or a functional fragment or derivative thereof.

All of these (poly)peptides may be advantageously used to enhance the solubility and/or prevent aggregation of the fusion proteins of the invention. Particularly preferred is to employ intein in said composition because the protein has been modified such that it undergoes a self-cleavage reaction at its N-terminus at low temperatures in the presence of thiols such as DTT (MPACT<sup>TM</sup> I System/New England Biolabs). Also comprised by this embodiment are functional fragments of

any of these above recited proteins. The term "functional fragment" as used herein is intended to denote the capability of said fragment to confer solubility or prevent aggregation.

Further preferred is that the nucleic acid contained in the composition of the invention is DNA. Particularly preferred is that said DNA is cDNA, synthetic DNA or (semi)synthetic DNA.

Also preferred is that the vector that may be contained in the composition of the invention is an expression vector or a gene targeting vector. These vectors may advantageously be used for transfecting hosts that may or may not be contained in the composition of the invention.

Such vectors may comprise further genes such as marker genes which allow for the selection of said vector in a suitable host cell and under suitable conditions. Preferably, the nucleic acid molecule of the invention is operatively linked to expression control sequences allowing expression in prokaryotic or eukaryotic cells. Expression of said nucleic acid molecule comprises transcription of the nucleic acid molecule into a translatable mRNA. Regulatory elements ensuring expression in eukaryotic cells, preferably mammalian cells, are well known to those skilled in the art. They usually comprise regulatory sequences ensuring initiation of transcription and optionally poly-A signals ensuring termination of transcription and stabilization of the transcript. Additional regulatory elements may include transcriptional as well as translational enhancers, and/or naturally-associated or heterologous promoter regions. Possible regulatory elements permitting expression in prokaryotic host cells comprise, e.g., the PL, lac, trp or tac promoter in *E. coli*, and examples for regulatory elements permitting expression in eukaryotic host cells are the AOX1 or GAL1 promoter in yeast or the CMV-, SV40-, RSV-promoter (Rous sarcoma virus), CMV-enhancer, SV40-enhancer or a globin intron in mammalian and other animal cells. Beside elements which are responsible for the initiation of transcription such regulatory elements may also comprise transcription termination signals, such as the SV40-poly-A site or the tk-poly-A site, downstream of the nucleic acid molecule. Furthermore, depending on the expression system used leader sequences capable of directing the polypeptide to a cellular compartment or secreting it into the medium may be added to the coding sequence of the nucleic acid molecule of the invention and are well known in the art. The leader sequence(s) is (are) assembled in appropriate phase with translation, initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein, or a portion thereof, into the periplasmic space or extracellular medium. Optionally, the

heterologous sequence can encode a fusion protein including an C- or N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. In this context, suitable expression vectors are known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3 (In-vitogene), or pSPORT1 (GIBCO BRL). Preferably, the expression control sequences will be eukaryotic promoter systems in vectors capable of transforming or transfecting eukaryotic host cells, but control sequences for prokaryotic hosts may also be used.

Gene therapy, which is based on introducing therapeutic genes into cells by ex-vivo or in-vivo techniques is one of the most important applications of gene transfer. Suitable vectors and methods for in-vitro or in-vivo gene therapy are described in the literature and are known to the person skilled in the art; see, e.g., Giordano, *Nature Medicine* 2 (1996), 534-539; Schaper, *Circ. Res.* 79 (1996), 911-919; Anderson, *Science* 256 (1992), 808-813; Isner, *Lancet* 348 (1996), 370-374; Muhlhäuser, *Circ. Res.* 77 (1995), 1077-1086; Wang, *Nature Medicine* 2 (1996), 714-716; WO94/29469; WO 97/00957 or Schaper, *Current Opinion in Biotechnology* 7 (1996), 635-640, and references cited therein. The nucleic acid molecules and vectors of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) into the cell. Preferably, said cell is a germ line cell, embryonic cell, or egg cell or derived therefrom, most preferably said cell is a stem cell.

Said hosts may then be used for the production of the fusion protein comprised in the composition of the invention.

Said host cell may be a prokaryotic or eukaryotic cell. The nucleic acid molecule or vector of the invention which is present in the host cell may either be integrated into the genome of the host cell or it may be maintained extrachromosomally.

Preferably, said host is a bacterial, preferably an *E.coli*, an animal-, preferably a mammalian, an insect-, a plant-, a fungal, preferably a yeast- and most preferably a *Saccharomyces* or *Aspergillus* cell, a *Pichia pastoris* cell, a transgenic animal or a transgenic plant.

The present invention also relates to a method of producing a fusion protein as defined in the diagnostic composition of any of the preceding claims comprising culturing or raising the host as defined in claim 11 and isolating said fusion protein. Preferably the bacterial host *E. coli* is used for the expression of the GST-HD fusion proteins. The cDNA fragments containing CAG repeats in the normal or pathological range may, for example, be cloned into pGEX-5X-1 (Pharmacia) and the resulting

plasmids expressing fusion proteins with polyglutamine-sequences used for protein purification. The resulting proteins may be purified under native conditions by affinity chromatography on glutathione agarose (Smith and Johnson, 1988).

Additionally, the present invention relates in a preferred embodiment to a composition wherein said antibody is a monoclonal antibody, polyclonal antibody, phage display antibody or a fragment or derivative thereof.

The above recited fragments or derivatives comprise Fv-, Fab-, F(ab)<sub>2</sub>-fragments, single-chain antibodies or single-chain antibody domains. These antibodies may advantageously be used in experiments such as Western blotting experiments to determine the presence of the protein on, for example, nitrocellulose membrane.

Alternatively, the antibody, derivative or fragment thereof may be used in immunometric assays such as ELISAs or RIAs or may be coupled to columns in order to retain the fusion protein from, for example, mammalian sources on said column for further detection or purification.

Alternative uses and advantages of the antibody or fragment or derivative contained in the composition of the invention are, on the basis of the teachings of this invention, clear to the person skilled in the art.

The invention further relates to an *in vitro* method of producing amyloid aggregates comprising (a) at least partially cleaving the fusion protein comprised in the diagnostic composition of the invention wherein the (poly)peptide that is released has the ability to self-assemble into amyloid-like fibrils or protein aggregates or (b) inducing self-assembly into amyloid-like fibrils or protein aggregates by heating the fusion protein comprised in the composition of the present invention or an amyloidogenic (poly)peptide that has the ability to self-assemble into amyloid-like fibrils or protein aggregates by inducing a pH change in a solution comprising said fusion protein/(poly)peptide or by treating said fusion protein/(poly)peptide with a denaturing agent.

The method of the invention may advantageously be used to study in more detail the process that leads to the formation of amyloid-like fibrils or protein aggregates from amyloidogenic (poly)peptides. Using the method of the invention, the onset of, for example, HD or AD can be examined in an *in vitro* situation. It is

important that the polypeptide that is released by the cleaving agent still retains the possibility to form amyloid-like fibrils or protein aggregates. The formation of such fibrils or aggregates may be monitored, for example, by electron or light microscopy. Varying the cleaving conditions in cases where more than one cleavage site is present on the fusion protein may be used to further elaborate the minimal requirements for said amyloid-like fibril or protein aggregate formation.

Preferably, the cleavage is effected chemically or enzymatically, or by the intein self-cleavage reaction in the presence of thiols.

In the following, enzymatic cleavage will also be referred to as proteolytic cleavage. The enzymatic cleavage has the advantage that the cleavage reaction can be performed under almost physiological conditions and normally only low amounts of the protease are necessary for the cleavage reaction. Furthermore, said cleavage is highly specific and the enzyme can be regarded as nontoxic. Therefore, one can envisage a wide variety of applications within this invention. The disadvantage of the enzymatic cleavage reaction is that prospective inhibitors might inhibit in some cases the protease and in turn prevent the formation of protein aggregates. In comparison, this is not the case when the cleavage reaction is performed chemically.

Further, the present invention relates to a method of testing a prospective inhibitor of aggregate formation of a fusion protein as defined in the composition of the invention when enzymatically or chemically cleaved or a non-cleaved fusion amyloidogenic (poly)peptide as defined hereinbefore or an amyloidogenic non-fusion (poly)peptide comprising

- (a) incubating in the presence of a prospective inhibitor
- (aa) said fusion protein in the presence or absence of a cleaving agent; or
- (ab) said non-fusion poly(peptide); and
- (b) assessing the formation of amyloid-like fibrils or protein aggregates.

This method of the present invention provides a particularly strong impact on the pharmaceutical research related to amyloid-associated diseases. For the first time, an inhibitor of fibril or aggregate formation can conveniently, directly, easily and within a short time be tested *in vitro*. As has been detailed herein above, aggregate formation may be tested on cleavage products, on non-cleaved fusion proteins or on the above recited non-fusion proteins which have the capacity to aggregate when the temperature is raised, the pH is lowered or the protein is dissolved in urea and the

urea is slowly diluted out with a solvent. Additionally, the present invention does not exclude self-assembly under different conditions.

It was shown recently that acid-mediated denaturation of, e.g., transthyretin yields a conformational intermediate that can self-assemble into amyloid (Lai et al., 1996). Booth et al. demonstrated that heat denaturation of human lysozyme variants resulted in instability, unfolding, and amyloid fibrillogenesis.

Preferably, the incubation is effected in the presence of factor Xa, trypsin, endoproteinase Arg-C, endoproteinase Lys-C, proteinase K, thrombin or elastase at a temperature of preferably 25 to 37°C for 0,5 to 16 hours and the assessment of the formation of fibrils or aggregates in step (b) is preferably effected by a filter assay or by a thioflavine T (ThT) fluorescence assay, in which the fluorescence intensity reflects the degree of aggregation.

As regards the filter assay, a more detailed protocol thereof is explained in the European patent application entitled "Novel method of detecting amyloid-like fibrils or protein aggregates" filed on the same day with the European Patent Office and assigned to the same applicant which is explicitly incorporated herein by reference.

Additionally, the present invention relates to a method for identifying an inhibitor of aggregate formation of a fusion protein as defined in the invention prior to or after proteolytic or chemical cleavage or of a non-fusion amyloidogenic (poly)peptide as described herein above comprising

- (a) loading a surface or gel with said protein or an aggregate thereof;
- (b) incubating said surface or gel with a prospective inhibitor; and
- (c) assessing whether the presence of said prospective inhibitor avoids or reduces aggregate formation or further aggregate formation.

In accordance with this embodiment of the invention, proteolytic or chemical cleavage can be advantageously effected either prior or after the loading of the surface or gel leaving the investigator additional degrees of freedom in devising his experiments. The method of the invention is both useful for investigating the onset of aggregate or fibril formation or assessing the progression of such a process starting from the already existing aggregate or fibril. The latter embodiment is particularly useful in investigating treatment conditions for patients that are already affected by the disease at an early or medium stage thereof.

There is strong evidence that the formation of amyloids is a nucleation-dependent polymerization similar to protein crystallization or microtubule assembly. However, the deposition of a monomer onto a preexisting amyloid template is independent of the nucleation process. Thus, it will be very important to study the deposition of monomers onto a defined template under physiological conditions. With our *in vitro* system we should be able to monitor the deposition of radiolabelled polyglu-containing monomers.

Preferably, said surface employed in the method of the invention is a membrane. Preferably, the membrane should be cellulose acetate and should have a low binding capacity for soluble proteins.

The invention also relates to an inhibitor identifiable or identified by the method of the invention.

The various methods described herein above will give rise to the isolation of a number of inhibitors which are also comprised by the present invention. Once such an inhibitor is known, it is of course not necessary to identify it again by the method of the invention. Rather, said inhibitor can be produced by chemical or recombinant means. In the case that the inhibitor is of proteinaceous material, it is preferred to resynthesize a compound having the or most of the characteristics of said inhibitor by peptidomimetics.

Preferably, a number of compounds or compound classes are tested for their efficacy to inhibit amyloid-like fibril or protein aggregate formation. Said compounds comprise an antibody, 4'-Iodo-4'-deoxydoxorubicin (IDOX), pyronine Y, guanidine hydrochloride, urea, rifampicin and derivatives thereof, myristyltrimethylammonium bromide, hydroquinone, p-benzoquinone, 1,4-dihydroxynaphthalene, p-methoxyphenol,  $\alpha$ -tocopherol, ascorbic acid,  $\beta$ -carotene, anthracycline, doxorubicin, hexadecyl-N-methylpiperidinium, dodecyltrimethyl-ammonium, N-tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate, a (poly)peptide, glutamine or an oligoglutamine peptide. These compounds may be used in the inhibition or reversion of aggregate formation, for example, by formulation into a pharmaceutical composition for the treatment of any of the diseases cited herein.

The present invention also relates to a pharmaceutical composition comprising the inhibitor of the invention and a pharmaceutically acceptable carrier and/or diluent.

The pharmaceutical composition of the invention will find wide applicability in the medical field. Essentially all diseases associated with protein aggregate formation or amyloid-like fibril formation, in particular if they are associated with neuronal tissue or cells, may be effectively treated with the pharmaceutical composition of the invention.

Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Compositions comprising such carriers can be formulated by well known conventional methods. These pharmaceutical compositions can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g., by intravenous, intraperitoneal, subcutaneous, intramuscular, topical or intradermal administration. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. A typical dose can be, for example, in the range of 0.001 to 1000 µg (or of nucleic acid for expression or for inhibition of expression in this range); however, doses below or above this exemplary range are envisioned, especially considering the aforementioned factors. Generally, the regimen as a regular administration of the pharmaceutical composition should be in the range of 1 µg to 10 mg units per day. If the regimen is a continuous infusion, it should also be in the range of 1 µg to 10 mg units per kilogram of body weight per minute, respectively. Progress can be monitored by periodic assessment. Dosages will vary but a preferred dosage for intravenous administration of DNA is from approximately  $10^6$  to  $10^{12}$  copies of the DNA molecule. The compositions of the invention may be administered locally or systemically. Administration will generally be parenterally, e.g., intravenously; DNA may also be administered directly to the target site, e.g., by ballistic delivery to an internal or external target site or by catheter to a site in an artery.

The therapeutically useful compounds identified according to the method of the invention may be administered to a patient by any appropriate method for the particular compound, e.g., orally, intravenously, parenterally, transdermally, transmucosally, or by surgery or implantation (e.g., with the compound being in the form of a solid or semi-solid biologically compatible and resorbable matrix) at or near the site where the effect of the compound is desired.

Finally, the present invention relates to a transgenic mammal or plant comprising a nucleic acid molecule or vector as described in the invention. The transgenic mammal or plant would advantageously be used for the *in vivo* testing of the efficacy of the inhibitors referred to above. With the pharmaceutical composition of the invention the formation of protein aggregates in the brain or other tissues of a transgenic animal can be monitored quantitatively. Furthermore, *in vivo* studies relating to the onset or progress of the above recited diseases may be carried out.

The figures show:

**Figure 1****SDS-PAGE Analysis of Purified GST and GST-HD Fusion Proteins.**

(a) Aliquots (15 ml) of eluates from the glutathione agarose column were subjected to 12.5 % SDS-PAGE and analyzed by staining with Coomassie blue R. Lanes 1-6 contain GST, GST-HD20, -HD30, -HD83 and -HD122, respectively; lane M contains molecular mass standards. (b) Proteins were transferred to nitrocellulose and probed with anti-HD1 antibody. Arrows mark the origin of electrophoresis.

**Figure 2****Structure of GST-HD fusion proteins.**

The amino acid sequence corresponding to exon 1 of huntingtin is boxed. Arrows labeled Xa and T indicate cleavage sites for factor Xa and trypsin, respectively.

**Figure 3****Site-Specific Proteolysis of GST-HD Fusion Proteins with Trypsin and Factor Xa.**

Tryptic digestions were performed at 37°C for 3 (a) or 16 h (b). Native proteins and their cleavage products were subjected to 12.5% SDS-PAGE, blotted onto nitrocellulose membranes, and probed with anti-HD1 antibody. Arrows mark the origin of electrophoresis. (c) Purified fusion proteins and their factor Xa and trypsin cleavage products were analyzed using the filter retardation assay. The proteins retained by the cellulose acetate and nitrocellulose membranes were detected by incubation with the anti-HD1 antibody.

**Figure 4****Electron Micrographs of Native GST-HD Fusion Proteins and their Factor Xa and Trypsin Cleavage Products.**

Purified GST fusion proteins were protease treated, negatively stained with uranyl acetate and viewed by electron microscopy. The undigested GST-HD51 molecules appear as a homogeneous population of small, round particles (a).

Removal of the GST-tag with factor Xa results in the formation of amyloid-like fibrils and intermediate structures (b + c). After partial digestion (3 h) of GST-HD51 with trypsin, the ribbons are associated with terminal clots (d, arrow), whereas prolonged digestion (16 h) produces ribbons without attached clots (e). Removal of the GST-tag from GST-HD20 shows no evidence for the formation of defined structures (f).

**Figure 5****Birefringence of Protein Aggregates Formed by Proteolytic Cleavage of GST-HD51.**

The protein aggregates were stained with Congo red. (a) Bright field, 200x; (b) Polarized light, 200x; (c) Polarized light, 100x.

**Figure 6****Polyglutamine-Containing Protein Aggregates are Formed *in vivo*.**

(a) Western blot analysis, after separation by 10% SDS-PAGE, of the nuclear (N) and cytosolic (C) protein fractions prepared from brain and kidney of an R6/2 hemizygous transgenic mouse and a littermate control. Blots were probed with anti-HD1, anti-GAPDH and anti-Fos B antibodies as indicated. (b) Detection of HD exon 1 protein aggregates formed *in vivo* using the cellulose acetate filter assay. The membrane was immunostained using the anti-HD1 antibody. (c) Ultrastructure of a neuronal intranuclear inclusion (NII). The presence of a NII in a striatal neuron of a 17 month old R6/5 homozygous mouse is shown. The NII is indicated by the large arrow and the fibrillar amyloid-like structures within the NII are indicated by two small arrows. The scale bar is 250 nm.

**Figure 7****Proposed Mechanism for the Formation of Amyloid-like Fibrils by Proteolytic Cleavage of GST-HD51.**

GST-HD51 molecules are represented as follows: zigzag line, elongated polyglutamine repeat forming a stable hairpin with  $\beta$ -sheet structure; dotted line, N-terminal amino acids in the HD exon 1 protein containing the factor Xa cleavage site (arrow) (undefined structure); thin line, C-terminal amino acids in the HD exon 1 protein (undefined structure); shaded symbol, the dimeric globule-like form of GST.

Prior to cleavage, the HD exon 1 protein is tightly bound to the GST-tag preventing intermolecular interactions (1). Removal of the GST-tag with factor Xa renders the polyglutamine repeat accessible allowing the formation of fibrils as seen by EM (3). During cleavage, intermediate structures form through specific polyglutamine interactions before complete release of the GST-tag has occurred (2). These intermediates appear as clots under EM and are frequently seen at the terminals of growing fibrils.

#### Figure 8

Structure of GST-HD fusion proteins. The amino acids sequence corresponding to the N-terminal portion of huntingtin is boxed and the amino acids corresponding to the biotinylation site are underlined. Arrows labeled (Xa) and (T) indicate cleavage sites for factor Xa and trypsin, respectively.

#### Figure 9

Detection of polyglutamine-containing protein aggregates formed *in vitro* and in transfected COS-1 cells using the dot-blot filter retardation assay. (A) Purified GST-HD20DP and -HD51DP fusion proteins (250 ng) and their factor Xa and trypsin cleavage products were applied to the filter as indicated. The aggregated proteins retained by the cellulose acetate membrane were detected by incubation with the anti-HD1 antibody. (B) Scanning electron micrograph of aggregated GST-HD51DP trypsin cleavage products retained on the surface of the cellulose acetate membrane (Photo: Heinrich Lündsdorf, GBF Braunschweig, Germany). (C) Dot-blot filter retardation assay performed on the insoluble fraction isolated from transfected and non-transfected COS-1 cells. COS-1 cells were transiently transfected with the plasmids pTL1-CAG20, -CAG51 and CAG93 encoding huntingtin exon 1 proteins with 20 (HD20), 51 (HD51) and 93 (HD93) glutamines, respectively. The pellet fractions obtained after centrifugation of whole cell lysates were subjected to DNase1/trypsin digestion, boiled in 2% SDS, and portions of 1, 3 and 6  $\mu$ l were filtered through a cellulose acetate membrane. The aggregated huntingtin protein retained on the membrane was detected with the anti-HD1 antibody. NT, non-transfected cells.

#### Figure 10

Detection and quantification of aggregates formed *in vitro* from biotinylated GST-HD exon 1 fusion proteins. Various amounts of the fusion proteins GST-HD51DPBio and -

HD20DPBio were filtered through a cellulose acetate membrane after a 3-h incubation at 37°C in the presence or absence of trypsin as indicated. (A) Images of the retained protein aggregates, detected with streptavidin-AP conjugate using either a fluorescent (upper panel) or a chemiluminescent AP substrate (lower panel). (B) Quantification of signal intensities obtained for the GST-HD51DPBio dots seen in A.. Fluorescence and chemiluminescence values are arbitrary units generated by the Lumi-Imager F1 and LumiAnalyst™ software (Boehringer Mannheim).

### Figure 11

Detection (A) and quantification (B) of aggregates formed *in vitro* from biotinylated GST-HD exon 1 fusion proteins using the dot-blot and microtitre plate filter retardation assay. Various amounts of the fusion proteins GST-HD51DPBio and -HD20DPBio were filtered through the cellulose acetate membranes after a 3-h incubation at 37°C in the presence or absence of trypsin as indicated. The detection and quantification of the aggregates was as described in Fig. 3.

### Figure 12

Detection of neurofibrillar tangles (NFTs) and  $\beta$ -amyloids in brain extracts prepared from Alzheimer's disease patients and controls using the dot-blot filter retardation assay. The cellulose acetate membrane was probed with the polyclonal anti-Tau, the monoclonal anti- $\beta$ -amyloid, or the polyclonal anti-HD antibody. A1, A2, and A3: protein extracts prepared from cerebral cortex of Alzheimer's disease patients; C1, C2, and C3: protein extracts prepared from cerebral cortex of normal individuals. GST-HD51, fusion of glutathione S-transferase and huntingtin exon 1 containing 51 glutamines.

The examples illustrate the invention:

#### Example 1:

#### Purification of GST-HD fusion proteins containing expanded polyglutamines

Exon 1 of the HD gene was isolated from genomic phage clones, derived from the normal and expanded alleles of an HD patient (Sathasivam et al., 1997), and used for the expression of GST-HD fusion proteins in *E. coli*. DNA fragments containing CAG repeats in the normal (CAG)<sub>20-33</sub> and expanded (CAG)<sub>37-130</sub> range were cloned into pGEX-5X-1 (Pharmacia), and the resulting plasmids expressing fusion proteins with 20 (GST-HD20), 30 (-HD30), 51 (-HD51), 83 (-HD83)

and 122 (-HD122) glutamines, respectively, were used for protein purification. For plasmid construction lambda phage from stock 91974 (Sathasivam et al., 1997) were plated to give single plaques which were innoculated into 400 ml cultures of *E. coli* XL1-Blue MRF' (Stratagene) for DNA preparation. The DNA sequence encoding the N-terminal portion of huntingtin (exon 1), including the CAG repeats, was amplified by PCR using the following pair of primers: ES 25 (TGGGATCCGCATGGCGACCCTGGAAAAGCTGATGAAGG) corresponding to nt315-343 of the HD gene (HDCRG, 1993) and containing a BamHI site (underlined) and ES 26 (GGAGTCGACTCACGGTCGGTGCAGCGGCTCCTCAGC) corresponding to nt516-588 and containing a Sall site (underlined). Conditions for PCR were as described (Mangiarini et al. 1996). Due to instability of the CAG repeat during propagation in *E. coli*, DNA preparations from individual plaques yielded different sized PCR products. Fragments of ~ 320, 360, 480, and 590 bp were gel-purified, digested with BamHI and Sall and inserted into the BamHI-Sall site of the expression vector pGEX-5X-1 (Pharmacia), yielding pCAG30, pCAG51, pCAG83 and pCAG122, respectively. pCAG20, containing 20' repeats of CAG within the cloned HD exon 1 sequence, was similarly constructed from a phage genomic clone derived from a normal allele. All constructs were verified by sequencing. After induction with IPTG, the resulting proteins were purified under native conditions by affinity chromatography on glutathione agarose. Thus, *E. coli* SCS1 (Stratagene) carrying the pGEX expression plasmid of interest was grown to an OD<sub>600nm</sub> of 0.6 and induced with IPTG (1 mM) for 3.5 h as described in the manufacturer's protocol (Pharmacia). Cultures (200 ml) of induced bacteria were centrifuged at 4000 g for 20 min, and the resulting pellets were stored at -80°C. Cells were thawed on ice and resuspended in 5 ml of lysis buffer (50 mM sodium phosphate, 150 mM NaCl, 1 mM EDTA, pH 7.4) containing 0.5 mg/ml lysozyme. After 45 min at 0°C, cells were sonicated with two 30 sec-bursts. Octyl-β-D-glucopyranoside was then added to a final concentration of 0.1% and the resulting lysate was clarified by centrifugation at 30,000 g for 30 min at 4°C. Cleared lysates were incubated for 1 h at 4°C with 500 µl of a 1:1 slurry of glutathione-agarose beads (Sigma) that had been washed times and resuspended in lysis buffer. The beads were poured into a small column and washed extensively with lysis buffer containing 0.1% octyl-β-D-glucopyranoside. The bound fusion protein was eluted with 2 ml of 15 mM glutathione (reduced) in lysis buffer. Typical yields were 0.5-1 mg of purified GST-HD20, -HD30 and -HD51 proteins per 200 ml of bacterial culture; yields of GST-HD83 and -HD122 were much lower, less than 10% of that obtained with the shorter fusion proteins. Protein was determined by the Bio-Rad dye binding assay using bovine serum albumin as standard. SDS-PAGE of the purified GST-HD20, -HD30, -HD51, -HD83 and -HD122

proteins revealed major bands of 42, 45, 50, 65 and 75 kDa, respectively (Fig. 1a). These bands were also detected when the various protein fractions were subjected to immunoblot analysis using the affinity purified anti-huntingtin antibody HD1 (Fig. 1b, lanes 2-6). HD1 specifically detects the GST-HD fusion proteins on immunoblots, whereas the GST-tag alone is not recognized (Fig. 1b, lane 1). For immunoblotting a bacterial plasmid encoding HD1-His, a His<sub>6</sub>-tagged fusion protein containing residues 1-222 of huntingtin, was generated by inserting a PCR-amplified *IT-15* cDNA fragment into the pQE-32 vector (Qiagen). The fusion protein was expressed in *E. coli*, affinity-purified under denaturating conditions on Ni-NTA agarose, and injected into rabbits. The resulting immune serum was then affinity-purified against the antigen that had been immobilized on Ni-NTA agarose. The GAPDH- and Fos B-specific antisera have been described (Wanker et al., 1997; Davies et al., 1997).

Western blotting was performed as detailed (Towbin et al., 1979). The blots were incubated with 1:1000 dilutions of the indicated primary antibody, followed by an alkaline-phosphatase-conjugated secondary antibody. Color development was carried out with 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium as substrates (Promega).

All recombinant proteins migrated at a size corresponding nearly to that predicted from their amino acid sequence. Interestingly, an additional high molecular weight band which remains at the top of the gel, was consistently detected in the protein fractions with the longest polyglns (83 and 122 residues; Fig. 1a and b, lane 5 and 6). This band was most prominent on the immunoblots but was also clearly detectable in the Commassie stained gel. This immunoreactive material was often still present at the bottom of the loading slots, even after the samples had been boiled for 5 min in the presence of 2% SDS and 6 M urea prior to loading.

#### **Example 2:**

#### **Proteolytic cleavage of GST-HD fusion proteins containing expanded polyglns**

It has been shown previously that the solubility of certain proteins can be enhanced by the addition of the GST-tag (Smith and Johnson, 1988) and it was therefore of interest to determine whether the removal of the GST-tag by proteolytic cleavage would have an effect on the solubility of the polygln-containing fusion proteins. Potential factor Xa and trypsin cleavage sites within the GST-HD fusion proteins are shown in Fig. 2. Factor Xa cleaves between the GST-tag and the HD exon 1 protein whereas trypsin removes an additional 15 amino acids from the N-

terminus and a single proline from the C-terminus, both proteases leaving the polygln repeat intact. The GST-HD20, -HD30 and -HD51 proteins were digested with trypsin under conditions designed to remove the GST-tag from the fusion protein without it being totally degraded. After cleavage, proteins were denatured by boiling in the presence of 2% SDS and analyzed by SDS-PAGE and immunoblotting using the anti-HD1 antibody. GST-HD20 and -HD30 cleavage yielded products migrating in a 12.5 % gel at approximately 30 and 33 kDa, respectively. In contrast, cleavage of GST-HD51 resulted in the formation of two protein products migrating at approximately 37 and 60 kDa, and an additional weak immunoreactive band on the bottom of the loading slots was also detected (Fig. 3a). This high molecular weight band was more pronounced when GST-HD51 was digested with trypsin under conditions in which the GST-tag was totally degraded (Fig. 3b). However, with proteins GST-HD20 and -HD30 this longer exposure to trypsin produced the same cleavage products as the ones seen in Fig. 3a and the high molecular weight products were not observed. Similar results were obtained with factor Xa protease and endoproteinases Arg-C and Lys-C. As regards the proteolytic cleavages, the following protocols were carried out: The GST-HD fusion proteins purified as described above were dialysed against 40 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1 mM EDTA and 5% (v/v) glycerol to raise the pH prior to proteolytic cleavage. The proteins were then combined with bovine factor Xa (New England Biolabs) or modified trypsin (Boehringer Mannheim, sequencing grade) in dialysis buffer containing 2 mM CaCl<sub>2</sub> at an enzyme:substrate ratio of 1:10 (w/w) or 1:40 (w/w), respectively. Incubations with factor Xa were at 25°C for 16 h. Tryptic digestions were performed at 37°C for 3 or 16 h as indicated. Digestions were terminated by the addition of PMSF to 1 mM. The degree of proteolysis was determined by SDS-PAGE followed by staining with Coomassie blue or immunoblotting using anti-HD1 antibody.

We have developed a simple and sensitive filter assay to detect the formation of high molecular weight insoluble protein aggregates. This assay is based on the finding that the SDS-insoluble protein aggregates obtained by proteolytic cleavage of GST-HD51 are retained on a cellulose acetate filter, whereas the soluble cleavage products of GST-HD20 and GST-HD30 are not. Factor Xa or trypsin digestions of purified GST-HD fusion proteins (10 µg) were performed in a 20 µl reaction mixture as described above. Reactions were terminated by adjusting the mixture to 2% SDS and 50 mM DTT. After heating at 100°C for 5 min, aliquots (0.5 µl) were diluted into 200 µl of 0.1% SDS and filtered through a cellulose acetate membrane (Schleicher & Schuell, 0.2 µm pore size) using a BRL dot blot filtration unit. Filters were washed

with water, and the SDS-insoluble aggregates retained on the filter detected by incubation with the anti-HD1 antibody, followed by an anti-rabbit secondary antibody conjugated to alkaline phosphatase (Boehringer Mannheim). Fig. 3c shows immunoblots of cellulose acetate and nitrocellulose membranes to which the native GST-HD20, -HD30 and -HD51 proteins and their factor Xa and trypsin cleavage products have been applied. On the cellulose acetate filter, only the cleavage products of GST-HD51 were detected by the anti-HD1 antibody, indicating the formation of insoluble high molecular weight protein aggregates. In contrast, all the uncleaved GST-HD fusion proteins and their digestion products were detected on the nitrocellulose control filter. This assay was also used to detect huntingtin aggregates present in a nuclear fraction from the brain of an R6/2 hemizygous mouse and littermate control (see preparation of nuclei below).

**Example 3:****Huntingtin proteins containing expanded polyglutins in the pathological range aggregate to amyloid-like birefringent fibrils**

Electron microscopy of negatively stained GST-HD51 fractions showed oligomeric particles with diameters of 6 to 7 nm (Fig. 4a); no higher ordered aggregates were observed. For electron microscopic observation, the native or protease-digested GST-HD fusion proteins were adjusted to a final concentration of 50 µg/ml in 40 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1 mM EDTA and 5% glycerol. Samples were negatively stained with 1% uranyl acetate and viewed in a Philips CM100 EM. In contrast, protein fractions obtained by proteolytic cleavage of GST-HD51 showed numerous clusters of high molecular weight fibrils and ribbon-like structures (Fig. 4b, c, d and e), reminiscent of purified amyloids (Prusiner et al., 1983). The fibrils obtained after digestion with factor Xa showed a diameter of 10-12 nm and their length varied from 100 nm up to several micrometers (Fig. 4b and c). In the trypsin-treated samples ribbon-like structures formed by lateral aggregation of fibrils with a diameter of 7.7 nm were observed (Fig. 4d and e). After treatment with factor Xa or limited digestion with trypsin, clots of small particles were frequently detected on one or both ends of the fibrils (Fig. 4b, c and d). These clots of varying sizes and shapes were not seen when GST-HD51 was digested with trypsin under conditions in which the GST-tag is totally degraded (Fig. 4e), indicating that they contain GST. In strong contrast to GST-HD51, the GST-HD20 and -HD30 proteins did not show any tendency to form ordered high molecular weight structures, either with or without protease treatment (Fig. 4f).

The insoluble protein aggregates formed by proteolytic cleavage of GST-HD51 were isolated by centrifugation and stained with Congo red (Caputo et al., 1992) and examined under a light microscope. For light microscopy, peptide aggregates formed by trypsin digestion of purified GST-HD fusion proteins (50 µg in 100 µl of digestion buffer) were collected by centrifugation at 30,000 g for 1 h and resuspended in 10 µl of water. Samples were mixed with 0.1 volume of a 2% (w/v) aqueous Congo Red (Sigma) solution, placed on aminoalkylsilane-coated glass slides, and allowed to dry overnight under a coverslip. After removing the coverslip, excess Congo Red was removed by washing with 90% ethanol. Evaluation of the Congo Red staining by polarization microscopy was performed using a Zeiss Axiolab Pol microscope equipped with strain-free lenses and optimally aligned cross-polarizers. After staining, the protein aggregates on the glass slides were red, indicating that they had bound the dye (Fig. 5a), and when examined under polarized light a green color and birefringence were detected (Fig. 5b and c). These staining characteristics were similar to those observed for prions (Prusiner et al., 1983) and amyloids (Caputo et al., 1992).

**Example 4:****Huntingtin proteins containing expanded polyGlns form amyloid-like protein aggregates *in vivo***

To determine whether the amyloid-like protein aggregates formed by proteolytic cleavage of GST-HD51 *in vitro* are also present *in vivo*, nuclear protein fractions of brain and kidney were prepared from mice transgenic for the HD mutation (line R6/2) and littermate controls (Davies et al., 1997; Mangiarini et al., 1996). Nuclei from the brain or kidney of an R6/2 hemizygous mouse with a repeat expansion of (CAG)<sub>143</sub> (Mangiarini et al., 1996) at ten weeks of age and littermate control were prepared as follows. Whole brain samples (80 mg) in 400 ml of 0.25 M sucrose in buffer A (50 mM triethanolamine [pH 7.5], 25 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5 mM DTT, 0.5 mM PMSF) were homogenized using 15 strokes of a tight-fitting glass homogenizer. The homogenate was adjusted to a final concentration of 5 mM DTT, and centrifuged at 800 g for 15 min. The supernatant was re-centrifuged at 100,000 g for 1 h, and the supernatant from this centrifugation was taken as the cytosolic fraction (fraction C). The loose pellet from the first centrifugation was homogenized, diluted to 1.2 ml with 0.25 M sucrose/buffer A, and mixed with two volumes of 2.3 M sucrose/buffer A. The mixture was then layered on top of 0.6 ml 2.3 M sucrose/buffer A in a SW60 tube and centrifuged at 124,000 g for 1 h. The pellet was harvested with a spatula, resuspended in 200 µl of 0.25 M sucrose/buffer A and

again centrifuged at 800 g for 15 min. The entire procedure was carried out at 4 °C. The pelleted nuclei were resuspended to a density of ~ 1 x 10<sup>7</sup> nuclei/ml in 0.25 sucrose/buffer A (fraction N) and stored at -80 °C. Nuclei from mouse kidney were prepared in the same way. The protein extracts were analyzed by SDS-PAGE and Western blotting using the anti-HD1 antibody (Fig. 6a). Strikingly, this antibody detected a prominent high molecular weight band in the nuclear fraction (N) prepared from R6/2 transgenic brain, very similar to the high molecular weight band obtained by proteolytic cleavage of GST-HD51 (Fig. 3b). No such immunoreactive band was detected in the nuclear fraction of brain from the littermate control and it was also absent from the corresponding cytoplasmic fractions (C). A small amount of high molecular weight material was also detected in the nuclear fraction prepared from R6/2 transgenic kidney, but was again absent from the cytoplasmic fraction. The purity of the nuclear and cytoplasmic fractions was confirmed by Western blot analysis using the anti-Fos B and anti-GAPDH antibodies. Anti-Fos B detected the transcription factor mainly in the nuclear fraction, and the enzyme GAPDH was only seen in the cytoplasmic fraction, as expected. The Western blot results were reproduced using the cellulose acetate filter assay (Fig. 6b). Using this assay, a 10-20 fold higher amount of transgene protein was detected in the nuclear fraction isolated from brain material, compared to that prepared from kidney.

The formation of NIs has been shown to precede the neuronal dysfunction that forms the basis of the progressive neurological phenotype observed in the R6 transgenic lines (Davies et al., 1997). These NIs are immunoreactive for both huntingtin and ubiquitin antibodies and contain the transgene but not the endogenous huntingtin protein. Therefore, Western blot analysis using an anti-ubiquitin antibody was also performed showing the same pattern of immunoreactivity as had been observed with the anti-HD1 antibody (Fig. 6a), and indicating that the high molecular weight transgene protein present in the nuclear fraction is ubiquitinated (data not shown).

To examine whether the NIs containing the proteins huntingtin and ubiquitin (Davies et al., 1997) have a fibrous composition, an ultrastructural analysis was performed. Experimentally, a 17 month old R6/5 homozygous mouse ((CAG)<sub>128-155</sub>) (Mangiarini et al., 1996) was deeply anaesthetised with sodium pentobarbitone and then perfused through the left cardiac ventricle with 35-50 ml of 4 % paraformaldehyde and either 0.5 % glutaraldehyde in 0.1 M Millonig's phosphate buffer (pH 7.4). The brain was removed from the skull and placed in fresh fixative overnight at 4 °C. Coronal sections (50 - 200 µm) were cut on an Oxford Vibratome

(Lancer) and collected in serial order in 0.1 M phosphate buffer. After being osmicated (30 min in 1% OsO<sub>4</sub> in 0.1 M phosphate buffer) the sections were stained for 15 min in 0.1 % uranyl acetate in sodium acetate buffer at 4 °C, dehydrated in ethanol, cleared in propylene oxide and embedded in Araldite between two sheets of Melanex (ICI). Semi thin (1 µm) sections were cut with glass knives and stained with toluidine blue adjacent to thin sections cut with a diamond knife on a Reichert Ultracut ultramicrotome. The sections were collected on mesh grids coated with a thin formvar film, counterstained with lead citrate and viewed in a Jeol 1010 electron microscope. An electron micrograph of a NII from a 17 month old R6/5 homozygous mouse is shown in Fig. 6c. This NII (large arrow) contains high molecular weight fibrous structures which were clearly differentiated from the surrounding chromatin. The filaments were randomly oriented, 5-10 nm in diameter and often measured up to 250 nm in length (small arrows). These structures differ from those previously reported in the NIIs seen in hemizygous R6/2 mice which were far more granular in composition, with individual filamentous structures being more difficult to distinguish (Davies et al., 1997). R6/2 mice exhibit an earlier age of onset with a more rapid progression of the phenotype and do not survive beyond 13 weeks (Mangiarini et al., 1996). It is possible that the filamentous structures do not have time to form in the R6/2 mice.

**Example 5:****Construction of further plasmids, purification of corresponding GST fusion proteins and proteolytic cleavage of GST fusion proteins**

In a second set of experiments, a further number of plasmids was constructed. Standard protocols for DNA manipulations were followed (J. Sambrook, E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Plainview, NY, 1989). *IT-15* cDNA sequences (HDCRG, *Cell* 72, 971 (1993)) encoding the N-terminal portion of huntingtin, including the CAG repeats, were amplified by PCR using the oligonucleotides ES25 (5'-TGGGATCCGCATGGCG ACCCTGGAAAGCTGATGAAGG-3') and ES27 (3'-CTCCTCGAGCGGGCGG TGGCGGCTGTTGCTGCTGCTGCTG-5') as primers and the plasmids pCAG20 and pCAG51 as template (E. Scherzinger, R. Lurz, M. Trumaine, L. Margiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker, *Cell* 90, 549 (1997)). Conditions for PCR were as described (L. Mangiarini, K. Sathasivam, M. Seller, B. Cozens, A. Harper, C. Hetherington, M. Lawton, Y. Trottier, H. Lehrach, S.W. Davies, and G.P. Bates, *Cell* 87, 493 (1996)). The resulting cDNA fragments were gel purified, digested with *Bam* HI and *Xho* I and were inserted into the

*Bam* HI-*Xho* I site of the expression vector pGEX-5X-1 (Pharmacia), yielding pCAG20DP and pCAG51DP, respectively. The plasmids pCAG20DP-Bio and pCAG51DP-Bio were generated by subcloning the PCR fragments obtained from the plasmids pCAG20 and pCAG51 into pGEX-5X-1-Bio. pGEX-5X-1-Bio was created by ligation of the oligonucleotides BIO1 (5'-CGCTCGAGGGTATCTTCGAGGCC) AGAAAGATCGAGTGGCGATCACCATGAG-3') and BIO2 (5'-GGCCGCTCATGGTG ATCGCCACTCGATCTTCTGGGCCTCGAAGATAACCTCGAG-3'), after annealing and digestion with *Xho* I, into the *Xho* I-*Not* I site of pGEX-5X-1. The plasmids with the *IT-15* cDNA inserts were sequenced to confirm that no errors had been introduced by PCR. The construction of plasmids pTL1-CAG20, pTL1-CAG51 and pTL1-CAG93 for the expression of huntingtin exon 1 proteins containing 20, 51 and 93 glutamines in mammalian cells has been described (A. Sittler, S. Wälter, N. Wedermeier, R. Hasenbank, E. Scherzinger, G. P. Bates, H. Lehrach, and E. E. Wanker, *Mol. Cell*, submitted).

The amino acid sequence of the GST-HD fusion proteins encoded by the *E. coli* expression plasmids pCAG20DP, pCAG51DP, pCAG20DP-Bio and pCAG51DP-Bio is shown in Fig. 8. The plasmids pCAG20DP and pCAG51DP encode fusion proteins of glutathione S-transferase (GST) and the N-terminal portion of huntingtin containing 20 (GST-HD20DP) and 51 (-HD51DP) polyglutamines, respectively. In these proteins the proline-rich region located immediately downstream of the glutamine repeat was deleted (E. Scherzinger, R. Lurz, M. Trumaine, L. Margiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker, *Cell* 90, 549 (1997)). The fusion proteins GST-HD20DPBio and -HD51DPBio are identical to GST-HD20DP and -HD51DP, except for the presence of a biotinylation site (P. J. Schatz, *Biotechnology* 11, 1138 (1993)) at their C-termini.

In the experiments described herein, *E. coli* DH10B (BRL) was used for plasmid construction and *E. coli* SCS1 (Stratagene) was used for the expression of GST-HD fusion proteins. Transformation of *E. coli* with plasmids and ligation mixtures was performed by electroporation using a Bio-Rad Gene Pulser (Richmond, CA). Transformed cells were spread on LB plates supplemented with appropriate antibiotics (J. Sambrook, E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Plainview, NY, 1989). For expression of GST fusion proteins, cells were grown in liquid TY medium (5 g NaCl, 5 g yeast extract, and 10 g tryptone per liter) buffered with 20 mM MOPS/KOH (pH 7.9) and supplemented with glucose (0.2%), thiamine (20 µg/ml), ampicillin (100 µg/ml) and kanamycin (25 µg/ml).

The procedure for purification of GST fusion proteins is an adaption of the protocol of Smith and Johnson (D. B. Smith and K. S. Johnson, *Gene* 67, 31 (1988)). Unless indicated otherwise, all steps were performed at 0-4°C.

First, 100 ml TY medium were inoculated with a single colony containing the expression plasmid of interest, and the culture was incubated at 37°C overnight with shaking. Then, 1.5 liter TY medium were inoculated with the overnight culture and grown at 37°C until an OD<sub>600</sub> of 0.6 was reached. IPTG was added to a final concentration of 1 mM, and the culture continued to grow at 37°C for 3.5 h with vigorous shaking. The culture was chilled on ice, and the cells harvested by centrifugation at 4000 x g for 20 min.

Cells were washed with buffer A [50 mM sodium phosphate (pH 8), 150 mM NaCl, and 1 mM EDTA]. If necessary, the cell pellet was stored at -70°C. Cells were resuspended in 25 ml buffer A. PMSF and lysozyme (Boehringer Mannheim) were added to 1 mM and 0.5 mg/ml, respectively, and incubated on ice for 45 min. Cells were lysed by sonication (2 x 45 s, 1 min cooling, 200-300 Watt), and Triton X-100 was added to a final concentration of 0.1% (v/v). The lysate was centrifuged at 30.000 x g for 30 min, and the supernatant was collected.

5 ml of a 1:1 slurry of GST-agarose (Sigma), previously equilibrated in buffer A, was added and the mixture was stirred for 30 min. The slurry was poured into a 1.6 cm diameter column, washed once with 40 ml buffer A containing 1 mM PMSF and 0.1 % Triton X-100 and twice with 40 ml buffer A containing 1 mM PMSF. The protein was eluted with 5 x 2 ml buffer A containing 15 mM reduced glutathione (Sigma). Aliquots of the fractions were analyzed by SDS-PAGE and the fractions containing purified GST fusion protein were combined. Finally, the pooled fractions were dialysed overnight against buffer B [20 mM Tris/HCl (pH 8), 150 mM NaCl, 0.1 mM EDTA and 5 % (v/v) glycerol], aliquotted, freezed in liquid nitrogen and stored at -70°C.

Typical yields were 10 - 20 mg for GST-HD20DP and -HD51DP and 5-10 mg for GST-HD20DPBio and -HD51DPBio per liter of bacterial culture. Protein concentration was determined using the Coomassie protein assay reagent from Pierce with BSA as a standard.

The GST-huntingtin fusion proteins (2 mg) were digested with bovine factor Xa (New England Biolabs) or with modified trypsin (Boehringer Mannheim, sequencing grade) at an enzyme/substrate ratio of 1:10 (w/w) and 1:20 (w/w), respectively. The reaction was

carried out in 20  $\mu$ l of 20 mM Tris/HCl (pH 8), 150 mM NaCl and 2 mM CaCl<sub>2</sub>. Incubations with factor Xa were performed at 25°C for 16 h. Tryptic digestions were at 37°C for 3 to 16 h. Digestions were terminated by the addition of 20  $\mu$ l 4% (w/v) SDS and 100 mM DTT, followed by heating at 98°C for 5 min.

As shown in the previous examples, removal of the GST tag from the HD exon 1 protein containing 51 glutamines (GST-HD51) by site-specific proteolytic cleavage results in the formation of high molecular weight protein aggregates, seen as characteristic fibrils or filaments on electron microscopic examination. Such ordered fibrillar structures were not detected after proteolysis of fusion proteins containing only 20 (GST-HD20) or 30 (GST-HD30) glutamines, although light scattering measurements (Y. Georgalis, E.B. Starikov, B. Hollenbach, R. Lurz, E. Scherzinger, W. Saenger, H. Lehrach, and E.E. Wanker, *Proc. Natl. Acad. Sci. USA* 95, 6118 (1998)) revealed that some form of aggregation also occurred with these normal repeat-length proteins. In the present example, truncated GST-HD exon 1 fusion proteins with or without a C-terminal biotinylation tag (P. J. Schatz, *Biotechnology* 11, 1138 (1993) were used. These fusion proteins contain either 20 or 51 glutamines but lack most of the proline rich region located downstream of the glutamine repeat (E. Scherzinger, R. Lurz, M. Trumaine, L. Margiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker, *Cell* 90, 549 (1997)). Potential factor Xa and trypsin cleavage sites within the GST-HD fusion proteins are shown in Fig. 8. As outlined above, the proteins GST-HD20DP and -HD51DP were expressed in *E. coli* and affinity-purified under native conditions. They were then digested overnight with trypsin or factor Xa protease to promote the formation of polyglutamine-containing huntingtin aggregates. Fig. 9A shows an immunoblot of a cellulose acetate membrane to which the native GST-HD20DP and -HD51DP proteins and their factor Xa and trypsin cleavage products have been applied.

To monitor the in vitro formation of polyglutamine-containing aggregates without the need for a specific antibody, a modified filter retardation assay was developed. In this assay, streptavidin-conjugated alkaline phosphatase (AP) is used to detect the insoluble protein aggregates retained on the cellulose acetate filter membrane. Streptavidin binds specifically to the biotinylation tag (P. J. Schatz, *Biotechnology* 11, 1138 (1993)) that has been added C-terminal to the polyglutamine tract in the fusion proteins GST-HD20DPBio and -HD51DPBio (Fig. 7) (see Example 8 for details). Fig. 10A shows that the modified aggregation assay gives results comparable to those obtained with the non-biotinylated fusion proteins in that insoluble aggregates are produced from the trypsin-treated GST-HD51DPBio protein but not from the uncleaved GST-HD51DPBio protein or the corresponding 20 repeat samples. Using either fluorescent (AttoPhos<sup>TM</sup>) or

chemiluminescent (CDP-Star™) substrates for alkaline phosphatase, it is possible to capture and quantify the filter assay results with the Boehringer Lumi-Imager F1 system. With both AP substrates, aggregates formed from as little as 5-10 ng of input GST-HD51DPBio protein were readily detected on the cellulose acetate membrane, and signal intensities increased linearly up to 250 ng of fusion protein applied to the filter (Fig. 10B).

**Example 6:****Isolation of amyloid-like protein aggregates from transfected COS-1 cells**

To examine whether polyglutamine-containing aggregates are also formed *in vivo*, HD exon 1 proteins with 20, 51 or 93 glutamines (without a GST tag) were expressed in COS-1 cells. Whole cell lysates were prepared, and after centrifugation, the insoluble material was collected and treated with DNaseI and trypsin to lower the viscosity. The resulting protein mixture was then boiled in SDS and analyzed using the dot-blot filter retardation assay (see Example 8). In more detail, the following experimental protocol was carried out:

COS-1 cells were grown in Dulbecco's modified Eagle medium (Gibco BRL) supplemented with 5% (w/v) fetal calf serum (FCS) containing penicillin (5 U/ml) and streptomycin (5 µg/ml), and transfection was performed as described (A. Sittler, D. Devys, C. Weber, and J.-L. Mandel, *Hum. Mol. Genet.* 5, 95 (1996)).

COS-1 cells transfected with the mammalian expression plasmids pTL1-CAG20, pTL1-CAG51 and pTL1-CAG93 were harvested 48 h after transfection. The cells were washed in ice cold PBS, scraped and pelleted by centrifugation (2000 x g, 10 min, 4°C). Cells were lysed on ice for 30 min in 500 ml lysis buffer [50 mM Tris/HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.5% (w/v) NP-40, 1 mM EDTA] containing the protease inhibitors PMSF (2 mM), leupeptin (10 µl/ml), pepstatin (10 µg/ml), aprotinin (1 µg/ml) and antipain (50 µg/ml). Insoluble material was removed by centrifugation for 5 min at 14000 rpm in a microfuge at 4°C. Pellets containing the insoluble material were resuspended in 100 ml DNase buffer [20 mM Tris/HCl (pH 8.0), 15 mM MgCl<sub>2</sub>], and DNase I (Boehringer Mannheim) was added to a final concentration of 0.5 mg/ml followed by incubation at 37°C for 1 h. After DNase treatment the protein concentration was determined by the Dot Metric assay (Geno Technology) using BSA as a standard. Eight µl 1 M Tris/HCl (pH 8.4), 1 µl 1% (w/v) SDS, 1 µl 0.2 M CaCl<sub>2</sub> and 10 µl trypsin (0.25 mg/ml) were then added, and the mixture was incubated for an additional 4 h at 37°C. Digestions were terminated by adjusting the mixtures to 20 mM EDTA, 2% (w/v) SDS and 50 mM DTT, followed by heating at 98°C for 5 min.

Fig. 9C shows that insoluble protein aggregates are being formed in transfected COS cells expressing the HD exon 1 protein with 51 and 93 glutamines but not in COS cells expressing the normal exon 1 allele with 20 glutamines or in the non-transfected control cells. Thus, as observed *in vitro* with purified GST fusion proteins, formation of high molecular weight protein aggregates *in vivo* occurs in a repeat length-dependent way and requires a polyglutamine repeat in the pathological range. In addition, like the *in vitro* aggregates, the HD exon 1 aggregates formed *in vivo* are resistant to digestion with trypsin as well as to boiling in 2% (w/v) SDS.

**Example 7:****Isolation of amyloid-like protein aggregates from Alzheimer's disease brain**

It has been shown that the neurodegenerative disorder Alzheimer's disease (AD) is caused by the formation of  $\beta$ -amyloids and neurofibrillar tangles (NFTs) mainly occurring in the neocortex, hippocampus and amygdala (K. Beyreuther, and C.L. Masters, *Nature* 383, 476 (1996)). To determine whether these structures can be detected by the dot-blot filter retardation assay brain extracts of patients and controls were prepared and analyzed using the anti-Tau, anti- $\beta$ -amyloid and anti-HD1 antibodies.

Fig. 12 shows that with the anti-Tau and anti- $\beta$ -amyloid antibodies NFTs and  $\beta$ -amyloids were detected in brain extracts prepared from patients A2 and A3, but not in brain extracts prepared from patient A1 and the controls. Clinical studies revealed that the patients A2 and A3 had Alzheimer's disease with an intermediate and severe intellectual impairment, respectively, whereas patient A1 suffered only from moderate intellectual impairment. This indicates that the results obtained with the filter retardation assay correlate with the severity of the disease. Using the HD1 antibody in the brain extracts prepared from AD patients and controls no aggregated huntingtin protein was detected. However, the antibody reacted with the GST-HD51 protein which was used as a positive control.

Human cerebral cortex (~ 500 mg) was homogenized in 2.5 ml of lysis buffer (0.32 M sucrose, 1 mM MgCl<sub>2</sub>, 5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 1 mM PMSF) using nine strokes of a glass homogenizer. The homogenate was centrifuged for 15 min at 500 x g to remove the nuclei. The original supernatant was then centrifuged at 93500 x g for 1 h yielding a membrane pellet. The pellet was dissolved in 2 - 5 ml 100 mM Tris-HCl (pH 8), 0.5% SDS and trypsin (Boehringer Mannheim, sequencing grade) was added to a final concentration of 0.05 mg/ml followed by incubation at 37°C overnight. Digestions were

terminated by adjusting the mixtures to 2% SDS and 50 mM DTT, followed by heating at 98°C for 5 min. The mixture was centrifuged for 1 h at 110000 x g and the resulting pellet was resuspended in 100 µl of water. Aliquots (2-10 µl) were then used for the analysis with the dot-blot filter retardation assay.

**Example 8:**

**Dot-blot filter retardation assay**

The filter assay used to detect polyglutamine-containing huntingtin protein aggregates has been described (hereinabove and in E. Scherzinger, R. Lurz, M. Trumaine, L. Margiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker, *Cell* 90, 549 (1997)). Denatured and reduced protein samples were prepared as described above, and aliquots corresponding to 50-250 ng fusion protein (GST-HD20DP and GST-HD51DP) or 5-30 µg extract protein (pellet fraction) were diluted into 200 µl 0.1% SDS and filtered on a BRL dot blot filtration unit through a cellulose acetate membrane (Schleicher and Schuell, 0.2 µm pore size) that had been preequilibrated with 0.1% SDS. Filters were washed 2 times with 200 µl 0.1% SDS and were then blocked in TBS (100 mM Tris/HCl, pH 7.4, 150 mM NaCl) containing 3% nonfat dried milk, followed by incubation with the anti-HD1 (1:1000) (see above and E. Scherzinger, R. Lurz, M. Trumaine, L. Margiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker, *Cell* 90, 549 (1997), the anti-Tau (Dako, 1:1000) or the anti-β-amyloid antibody (Dako, 1:300). The filters were washed several times in TBS, then incubated with a secondary anti-rabbit or anti-mouse antibody conjugated to horse radish peroxidase (Sigma, 1:5000) followed by ECL (Amersham) detection. The developed blots were exposed for various times to Kodak X-OMAT film or to a Lumi-Imager (Boehringer Mannheim) to enable quantification of the immunoblots.

For detection and quantification of polyglutamine-containing aggregates generated from the protease-treated fusion proteins GST-HD20DPBio and -HD51DPBio, the biotin/streptavidin-AP detection system was used. Following filtration, the cellulose acetate membranes were incubated with 1% (w/v) BSA in TBS for 1 h at room temperature with gentle agitation on a reciprocal shaker. Membranes were then incubated for 30 min with streptavidin-alkaline phosphatase (Promega) at a 1:1000 dilution in TBS containing 1% BSA, washed 3 times in TBS containing 0.1% (v/v) Tween 20 and 3 times in TBS, and finally incubated for 3 min with either the fluorescent alkaline phosphatase substrate AttoPhos™ or the chloro-substituted 1,2-dioxetane chemiluminescence substrate CDP-Star™ (Boehringer Mannheim) in 100 mM Tris/HCl, pH 9.0, 100 mM NaCl and 1 mM MgCl<sub>2</sub>. Fluorescent and chemiluminescent signals were

imaged and quantified with the Boehringer Lumi-Imager F1 system and LumiAnalyst<sup>TM</sup> software (Boehringer Mannheim).

**Example 9:**

**Microtitre plate filter retardation assay**

To process a large number of proteolytic digestion reactions in parallel, a microtitre plate filter retardation assay was developed. In this assay a 96-well microtitre plate containing a cellulose acetate membrane with a pore size of 0.45 mm (Whatman Polyfiltrronics) was used for the retention of polyglutamine-containing protein aggregates.

The following experimental protocol was employed:

First, 15 µl GST fusion protein solution (200 µg/ml GST-HD51DPBio or GST-HD20DPBio in buffer P [20 mM Tris/HCl (pH 8.0), 150 mM NaCl]) and 15 µl trypsin solution (10 µg/ml trypsin (Boehringer Mannheim, sequencing grade) in buffer P) were combined in a 96-well Thermo-Fast@96 tube plate (Advanced Biotechnologies LTD) using a multi channel pipette (Eppendorf), and the microtitre plate was incubated for 16 hours at 37°C. Then 30 µl SDS/DTT solution (4% SDS, 100 mM DTT in buffer P) were added to each well, the plate was sealed with a microtitre plate sealer (Biostat LTD) and the plate was heated in a 96-well MasterCycler (Eppendorf-Netheler-Hinz) for 5 min at 98°C.

The sealing was removed and 50 µl of the reaction mix were transferred into each well of a new 96-well microtitre plate containing a 0.45 µm cellulose acetate membrane, pre-equilibrated with 0.1% (w/v) SDS, using a multi channel pipette. For equilibration of the cellulose acetate membrane, the microtitre plate was placed into the QIAvac Manifold-96 (Qiagen) and 200 µl 0.1% SDS was pipetted into each well of the plate. Vacuum was then applied until the SDS solution had passed through the filter. Prior to addition of the protein solution, each well of the filter plate was preloaded with an additional 200 µl of 0.1% SDS. The diluted protein solution was then filtered through the membrane by applying vacuum.

The filterplate was washed with 2 x 200 µl 0.1% SDS and 2 x 200 ml TBS (100 mM Tris/HCl (pH 7.4), 150 mM NaCl). Vacuum was used to remove wash solutions from the membrane. 200 µl 0.2% (w/v) BSA in TBS were pipetted into each well of the filterplate, and the plate was incubated for 1 h at room temperature (RT) (blocking). Blocking buffer was removed by pipetting.

Next, 200  $\mu$ l streptavidin alkaline phosphatase (1:1000, Promega) in 0.2% (w/v) BSA/TBS were added to each sample, and the filterplate was incubated for 1 h at RT. Streptavidin AP buffer was removed by pipetting. The filterplate was washed with 3 x 200  $\mu$ l TTBS [100 mM Tris/HCl (pH 7.4), 150 mM NaCl, 0.1% (v/v) Tween 20] and 3 x 200  $\mu$ l TBS. Vacuum was used to remove wash solutions.

200  $\mu$ l detection buffer (50 mM Tris/HCl (pH 9.0), 500 mM NaCl, 1 mM Mg Cl<sub>2</sub>) were added to each sample, incubated for 1 min and vacuum was applied to remove the buffer. 200  $\mu$ l Attophos<sup>TM</sup> (10 mM AttoPhos<sup>TM</sup>) in detection buffer were pipetted into each well of the filterplate, incubated for 1 h at RT, vacuum was applied to remove the buffer, and the fluorescence emission of each well was measured with the CytoFluor<sup>®</sup>4000 (Perceptive Biosystems) at 485+/-20 (excitation) and 530 +/-25 (emission). Finally, the resultant images were analysed with CytoFluor 4.1 software and MS Excel 7.0.

As expected from the text set of experiments, using fusions of GST and the full-length HD exon 1 protein, only the cleavage products of GST-HD51DP were retained by the filter and were detected by the huntingtin-specific antibody HD1, indicating the formation of high molecular weight HD51DP aggregates from this fusion protein. Scanning electron microscopy of the material retained on the surface of the membrane revealed bunches of long fibrils or filaments (Fig. 9B), which were not detected after filtration of the uncleaved GST-HD51DP preparation or the protease-treated GST-HD20DP preparation. These results indicate that an elongated polyglutamine sequence but not the proline rich region in the HD exon 1 protein is necessary for the formation of high molecular weight protein aggregates *in vitro*.

Essentially, the same results as with the dot blot-filter retardation assay were obtained when the fusion proteins GST-HD20DPBio and -HD51DPBio were analysed with the microtitre plate filter retardation assay, indicating that this assay can be used for the high throughput isolation of chemical compounds from chemical libraries (Fig. 11A and B).

## References

Bates, G. P., Mangiarini, L., Mahal, A. and Davies, S. W. (1997). Transgenic models of Huntington's disease. *Hum. Mol. Genet.* 6, 1633-1637.

Booth, D. R., Stunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* 385, 787-793.

Burke, J. R., Enghild, J. J., Martin, M. E., Jou, Y.-S., Myers, R. M., Roses, A. D., Vance, J. M., and Strittmatter, W. J. (1996). Huntington and DRPLA proteins selectively interact with the enzyme GAPDH. *Nature Med.* 2, 347-350.

Caputo, C. B., Fraser, P. E., Sobel, I. E., and Krischner, D. A. (1992). Amyloid-like properties of a synthetic peptide corresponding to the carboxy terminus of b-amyloid protein precursor. *Arch. Bioch. Biophys.* 292, 199-205.

Caughey, B., and Chesebro, B. (1997). Prion protein and the transmissible spongiform encephalopathies. *Trends Cell Biol.* 7, 56-62.

Davies, S. W., Trumaine, M., Cozens, B. A., DiFiglia, M., Sharp, A. H., Ross, C. A., Scherzinger, E., Wanker, E. E., Mangiarini, L., and Bates, G. P. (1997). Formation of neuronal intranuclear inclusions (NII) underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90, 537-548.

de Rooij, K. E., Dorsman, J. C., Smoor, M. A. T., d. D. J., and van Ommen, G.-J. (1996). Subcellular localisation of the Huntington's disease gene product in cell lines by immunofluorescence and biochemical subcellular fractionation. *Hum. Mol. Genet.* 5, 1093-1099.

DiFiglia, M., Sapp, E., Chase, K., Schwarz, C., Meloni, A., Young, C., Martin, E., Vonstattel, J.-P., Carraway, R., Reeves, S. A., Boyce, F. M., and Aronin, N. (1995). Huntington is a cytoplasmic protein associated with vesicles in human and rat brain neurons. *Neuron* 14, 1075-1081.

Duyao, M. P., Auerbach, A. A., Ryan, A., Persichetti, F., Barnes, G. T., McNeil, S. M., Ge, P., Vonstattel, J.-P., Gusella, J. F., Joyner, A. L., and MacDonald, M. E. (1995). Inactivation of the mouse Huntington's disease gene homolog *Hdh*. *Science* 269, 407-410.

Glenner, G. G. (1980). Amyloid deposits and amyloidosis. *N. Engl. J. Med.* 302, 1283-1292, 1333-1343.

Goldberg, Y. P., Nicholson, D. W., Rasper, D. M., Kalchman, M. A., Koide, H. B., Graham, R. K., Bromm, M., Kazemi-Esfarjani, P., Thomberry, N. A., Vaillancourt, J. P., and Hayden, M. R. (1996). Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract. *Nature Genet.* 13, 442-449.

Gutekunst, C.-A., Levey, A. I., Heilman, C. J., Whaley, W. L., Yi, H., Nash, N. R., Rees, H. D., Madden, J. J., and Hersch, S. M. (1995). Identification and localization of huntingtin in brain and human lymphoblastoid cell lines with anti-fusion protein antibodies. *Proc. Natl. Acad. Sci. USA* 92, 8710-8714.

Harper, P. S. (1991). Huntington's disease, 22 Edition, P. S. Harper, ed. (London: W.B. Saunders Co, Ltd).

HDCRG (1993). A novel gene containing a trinucleotide repeat that is unstable on Huntington's disease chromosomes. *Cell* 72, 971-983.

Hoogeveen, A. T., Willemsen, R., Meyer, N., de Rooij, K. E., Roos, R. A. C., van Ommen, G.-J. B., and Galjaard, H. (1993). Characterisation and localisation of the Huntington disease gene product. *Hum. Mol. Genet.* 2, 2069-2073.

Ikeda, H., Yamaguchi, M., Sugai, S., Aze, Y., Narumiya, S., and Kakizuka, A. (1996). Expanded polyglutamine in the Machado-Joseph disease protein induces cell death *in vitro* and *in vivo*. *Nature Genet.* 13, 196-202.

Jarrett, J. T., and Lansbury, P. T. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73, 1055-1058.

Kalchman, M. A., Graham, R. K., Xia, G., Koide, H. B., Hodgson, J. G., Graham, K. C., Goldberg, Y. P., Gietz, R. D., Pickart, C. M., and Hayden, M. R. (1996). Huntington is ubiquinated and interacts with a specific ubiquitin-conjugating enzyme. *J. Biol. Chem.* 271, 19385-19394.

Kalchman, M. A., Koide, H. B., McCutcheon, K., Graham, R. K., Nichol, K., Nishiyama, K., Kazemi-Esfariani, P., Lynn, F. C., Wellington, C., Metzler, M., Goldberg, Y. P., Kanazawa, I., Gietz, R. D., and Hayden, M. R. (1997). *HIP1*, a human homologue of *S. cerevisiae* Sla2p, interacts with membrane-associated huntingtin in the brain. *Nature Genet.* 16, 44-53.

Li, X.-J., Li, S.-H., Sharp, A. H., Nucifora, F. C., Schilling, G., Lanahan, A., Worley, P., Snyder, S. H., and Ross, C. A. (1995). A huntingtin-associated protein enriched in brain with implications for pathology. *Nature* 378, 398-402.

Lim, K., Ho, J. X., Keeling, K., Gilliland, G. L., Ji, X., Rüker, F., and Carter, D. C. (1994). Three-dimensional structure of *Schistosoma japonicum* glutamine S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. *Prot. Sci.* 3, 2233-2244.

Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996). Exon 1 of the Huntington's disease gene containing a highly expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87, 493-506.

Onodera, O., Roses, A. D., Tsuji, S., Vance, J. M., Stritmatter, W. J., and Burke, J. R. (1996). Toxicity of expanded polyglutamine-domain proteins in *Escherichia coli*. *FEBS Lett.* 399, 135-139.

Perutz, M. F. (1996). Glutamine repeats and inherited neurodegenerative diseases: molecular aspects. *Curr. Opin. Struct. Biol.* 6, 848-858.

Perutz, M. F., Johnston, T., Suzuki, M., and Finch, J. T. (1994). Glutamine repeats as polar zippers: their possible role in neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* 91, 5355-5358.

Portera-Cailliau, C., Hedreen, J. C., Price, D. L., and Koliatsos, V. E. (1995). Evidence of apoptotic cell death in Huntington disease and excitotoxic animal models. *J. Neurosci.* 15, 3775-3787.

Prusiner, S. B., Kinley, M. P. M., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. F., and Glenner, G. G. (1983). Scrapie prions aggregate to form amyloid-like birefringent rods. *Cell* 35, 349-358.

Roizin, L., Stellar, S., and Liu, J. C. (1979). Neuronal nuclear-cytoplasmic changes in Huntington's Chorea: Electron microscope investigations. *Adv. Neurol.* 23, 95-122.

Roos, R. A. C., and Bots, G. T. A. M. (1983). Nuclear membrane indentations in Huntington's chorea. *J. Neurol. Sci.* 61, 37-47.

Ross, C. A. (1995). When more is less: pathogenesis of glutamine repeat neurodegenerative diseases. *Neuron* 15, 493-496.

Rubinsztein, D. C., Leggo, J., Coles, R., Almqvist, E., Biancalana, V., Cassiman, J.-J., Chotai, K., Connarty, M., Crauford, D., Curtis, A., Curtis, D., Davidson, M. J., Differ, A.-M., Dode, C., Dodge, A., Frontali, M., Ranen, N. G., Stine, O. C., Sherr, M., Abbott, M. H., Franz, M. L., Graham, C. A., Harper, P. S., Hedreen, J. C., Jackson, A., Kaplan, J.-C., Losekoot, M., MacMillan, J. C., Morrison, P., Trottier, Y., Novelletto, A., Simpson, S. A., Theilmann, J., Whittaker, J. L., Folstein, S. E., Ross, C. A., and Hayden, M. R. (1996). Phenotypic characterisation of individuals with 30-40 CAG repeats in the Huntington's disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *Am. J. Hum. Genet.* 59, 16-22.

Sathasivam, K., Amaechi, I., Mangiarini, L., and Bates, G. P. (1997). Identification of an HD patient with a (CAG)180 repeat expansion and the propagation of highly expanded CAG repeats in lambda phage. *Hum. Genet.* 99, 692-695.

Sharp, A. H., Loev, S. J., Schilling, G., Li, S.-H., Li, X.-J., Bao, J., Wagster, M. V., Kotzuk, J. A., Steiner, J. P., Lo, A., Hedreen, J., Sisodia, S., Snyder, S. H., Dawson, T. M., Ryugo, D. K., and Ross, C. A. (1995). Widespread expression of Huntington's disease gene (IT15) protein product. *Neuron* 14, 1065-1074.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of peptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67, 31-40.

Stott, K., Blackburn, J. M., Butler, P. J. G., and Perutz, M. (1995). Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* 92, 6509-6513.

Tellez-Nagel, I., Johnson, B., and Terry, R. D. (1974). Studies on brain biopsies of patients with Huntington's chorea. *J. Neurocyt.* 3, 308-332.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76, 4350-4354.

Trottier, Y., Devys, D., Imbert, G., Sandou, F., An, I., Lutz, Y., Weber, C., Agid, Y., Hirsch, E. C., and Mandel, J.-L. (1995a). Cellular localisation of the Huntington's disease protein and discrimination of the normal and mutated forms. *Nature Genet.* 10, 104-110.

Trottier, Y., Lutz, Y., Stevanin, G., Imbert, G., Devys, D., Cancel, G., Sandou, F., Weber, C., David, G., Tora, L., Agid, Y., Brice, A., and Mandel, J.-L. (1995b). Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 378, 403-406.

Vonsattel, J.-P., Myers, R. H., Stevens, T. J., Ferrante, R. J., Bird, E. D., and Richardson, E. P. (1985). Neuropathological classification of Huntington's disease. *J. Neuropath. Exp. Neurol.* 44, 559-577.

Wanker, E. E., Rovira, C., Scherzinger, E., Hasenbank, R., Walter, S., Tait, D., Colicelli, J., and Lehrach, H. (1997). HIP-1: A huntingtin interacting protein isolated by the yeast two-hybrid system. *Hum. Mol. Genet.* 6, 487-495.

M. DiFiglia, E. Sapp, K. O. Chase, S. W. Davies, G. P. Bates, J. P. Vonsattel, and N. Aronin, *Science* 277, 1990 (1997).

M. W. Becher, J. A. Kotzuk, A. H. Sharp, S. W. Davies, G. P. Bates, D. L. Price, and C. A. Ross, *Neurobiol. Disease* 4, 1 (1997).

S. Igarashi, R. Koide, T. Shimohata, M. Yamada, Y. Hayashi, H. Takano, H. Date, M. Oayke, T. Sato, A. Sato, S. Egawa, T. Ikeuchi, H. Tanaka, R. Nakano, K. Tanaka, I. Hozumi, T. Inuzuka, H. Takahashi, and S. Tsuji, *Nature Genetics* 18, 11 (1998).

P. J. Skinner, B. T. Koshy, C. J. Cummings, I. A. Klement, K. Helin, A. Servadio, H. Y. Zoghbi, and H. T. Orr, *Nature* 389, 971 (1997).

A. Matilla, B. T. Koshy, C. J. Cummings, T. Isobe, H. T. Orr, and H. Y. Zoghbi, *Nature* 389, 974 (1997).

H. L. Paulson, M. K. Perez, Y. Trottier, J. Q. Trojanowski, S. H. Subramony, S. S. Das, P. Vig, J.-L. Mandel, K. H. Fischbeck, and R. N. Pittman, *Neuron* 19, 333 (1997).

M. Holmberg, C. Duyckaerts, A. Durr, G. Cancel, I. Gourfinkel-An, P. Damier, B. Faucheux, Y. Trottier, E. C. Hirsch, Y. Agid, and A. Brice, *Hum. Mol. Genet.* 7, 913 (1998).

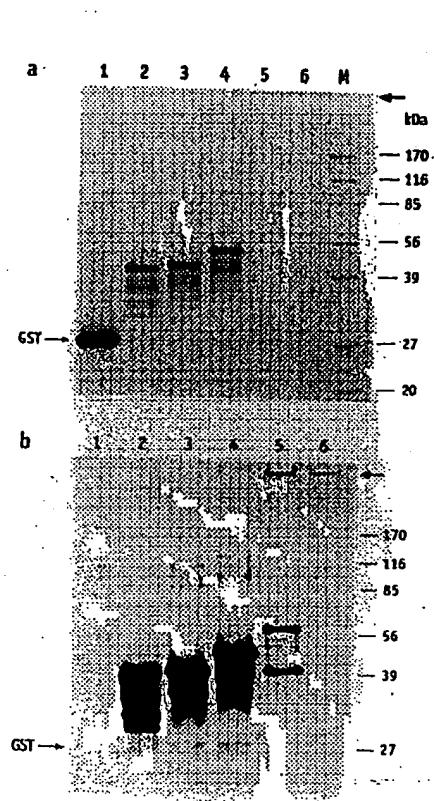
## CLAIMS

1. A composition comprising
  - (a) a nucleic acid molecule encoding a fusion protein comprising
    - (aa) a (poly)peptide that enhances solubility and/or prevents aggregation of said fusion protein; and
    - (ab) an amyloidogenic (poly)peptide that has the ability to self-assemble into amyloid-like fibrils or protein aggregates;
    - (b) a vector containing the nucleic acid molecule of (a);
    - (c) a host transformed with the vector of (b);
    - (d) a fusion protein encoded by the nucleic acid of (a) or a functional derivative thereof; and/or
    - (e) an antibody specific for the fusion protein of (d).
2. The composition of claim 1 wherein the amyloidogenic (poly)peptide comprises a polyglutamine expansion.
3. The composition of claim 2 wherein said polyglutamine expansion comprises at least 35 glutamines.
4. The composition of claim 3 wherein said polyglutamine expansion comprises at least 51 glutamines.
5. The composition of any one of claims 2 to 4 wherein said (poly)peptide defined in (ab) is huntingtin, androgen receptor, atropin, TATA binding protein, or ataxin-1,-2,-3, -6 or -7 or a fragment or derivative thereof.
6. The composition of any one of claims 1 to 5 wherein said amyloidogenic (poly)peptide self-assembles subsequent to release from said fusion protein.
7. The composition of claim 1 wherein said amyloidogenic (poly)peptide is the amyloid precursor protein (APP),  $\beta$ -protein, an immunoglobulin light chain, serum amyloid A, transthyretin, cystatin C,  $\beta$ 2-microglobulin, apolipoprotein A-1, gelsoline, islet amyloid polypeptide (IAPP), calcitonin, a prion, atrial natriuretic factor (ANF), lysozyme, insulin, fibrinogen, tau proteins or  $\alpha$ -synuclein or a fragment or derivative thereof.

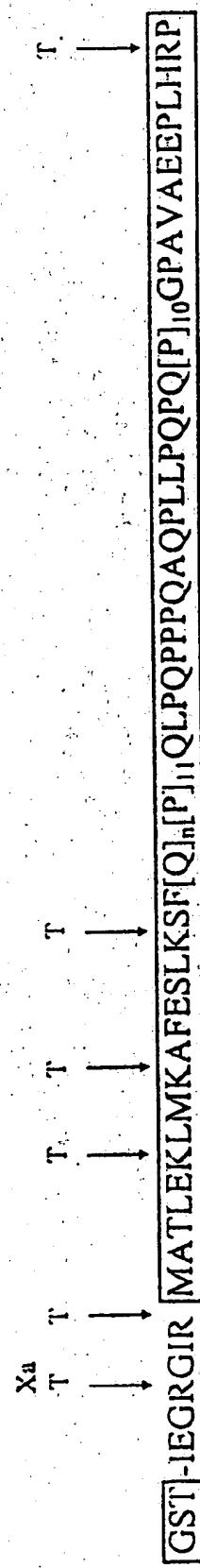
8. The composition of any one of claims 1 to 7 wherein said (poly)peptide defined in (aa) is glutathione S-transferase (GST), intein, thioredoxin, dihydrofolate reductase (DHFR) or chymotrypsin inhibitor 2 (CI2) or a functional fragment or derivative thereof.
9. The composition of any one of claims 1 to 8 wherein said nucleic acid is DNA.
10. The composition of any one of claim 1 to 9 wherein said vector is an expression vector or a gene targeting vector.
11. The composition of any one of claims 1 to 10 wherein said host is a bacterial, preferably an E.coli, an animal-, preferably a mammalian, an insect-, a plant-, a fungal, preferably a yeast- and most preferably a *Saccharomyces* or *Aspergillus* cell, a *Pichia pastoris* cell, a transgenic animal or a transgenic plant.
12. A method of producing a fusion protein as defined in the composition of any of the preceding claims comprising culturing or raising the host as defined in claim 11 and isolating said fusion protein.
13. The composition of any one of claims 1 to 12 wherein said antibody is a monoclonal antibody, polyclonal antibody, phage display antibody or a fragment or derivative thereof.
14. An in vitro method of producing amyloid aggregates comprising
  - (a) at least partially cleaving the fusion protein comprised in the composition of any one of claims 1 to 13 wherein the (poly)peptide that is released has the ability to self-assemble into amyloid-like fibrils or protein aggregates; or
  - (b) inducing self-assembly into amyloid-like fibrils or protein aggregates by heating the fusion protein comprised in the composition of any one of claims 1 to 13 or an amyloidogenic (poly)peptide that has the ability to self-assemble into amyloid-like fibrils or protein aggregates, by inducing a pH change in a solution comprising said fusion protein/(poly)peptide or by treating said fusion protein/(poly)peptide with a denaturing agent.
15. The method of claim 14 wherein said cleavage is effected chemically or enzymatically, or by the intein self-cleavage reaction in the presence of thiols.

16. A method of testing a prospective inhibitor of aggregate formation of a fusion protein as defined in the composition of any one of claims 1 to 13 when enzymatically or chemically cleaved or a non-cleaved fusion amyloidogenic (poly)peptide as defined in any one of claims 1 to 13 or an amyloidogenic non-fusion (poly)peptide comprising
  - (a) incubating in the presence of a prospective inhibitor
  - (aa) said fusion protein in the presence or absence of a cleaving agent; or
  - (ab) said non-fusion poly(peptide); and
  - (b) assessing the formation of amyloid-like fibrils or protein aggregates.
17. The method of claim 16 wherein incubation is effected in the presence of factor Xa, trypsin, endoproteinase Arg-C, endoproteinase Lys-C, proteinase K or elastase at a temperature of preferably 25 to 37°C for 0.5 to 16 hours and the assessment of the formation of fibrils or aggregates in step (b) is preferably effected by a filter assay or by a thioflavine T (ThT) fluorescence assay, in which the fluorescence intensity reflects the degree of aggregation.
18. A method for identifying an inhibitor of aggregate formation of a fusion protein as defined in any one of claims 2 to 6 prior to or after proteolytic or chemical cleavage or of a non-fusion amyloidogenic (poly)peptide that has the ability to self-assemble into amyloid-like fibrils or protein aggregates comprising
  - (a) loading a surface or gel with said protein or an aggregate thereof;
  - (b) incubating said surface or gel with a prospective inhibitor; and
  - (c) assessing whether the presence of said prospective inhibitor avoids or reduces aggregate formation or further aggregate formation.
19. The method of claim 18 wherein said surface is a membrane.
20. An inhibitor identifiable or identified by the method of any one of claims 16 to 19 or a derivative thereof.
21. The inhibitor of claim 20 which is an antibody, 4'-Iodo-4'-deoxydoxorubicin (IDOX), pyronine Y, guanidine hydrochloride, urea, rifampicin or a derivative thereof, myristyltrimethylammonium bromide, hydroquinone, p-benzoquinone, 1,4-dihydroxynaphthalene, p-methoxyphenol,  $\alpha$ -tocopherol, ascorbic acid,  $\beta$ -carotene, anthracycline, doxorubicin, hexadecyl-N-methylpiperidinium, dodecyltrimethyl-ammonium, N-tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate, a (poly)peptide, glutamine or an oligoglutamine peptide.

22. A pharmaceutical composition comprising the inhibitor of claim 20 or 21 and a pharmaceutically acceptable carrier and/or diluent.
23. A transgenic mammal or plant comprising a nucleic acid molecule or vector as described in the composition of claim 3 or 4.



Figur 1  
1/12



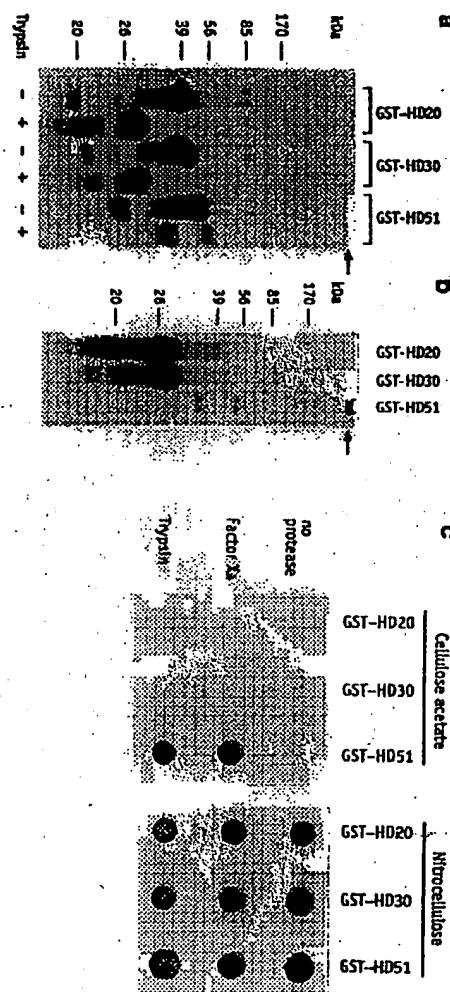


Figure 3  
3/12

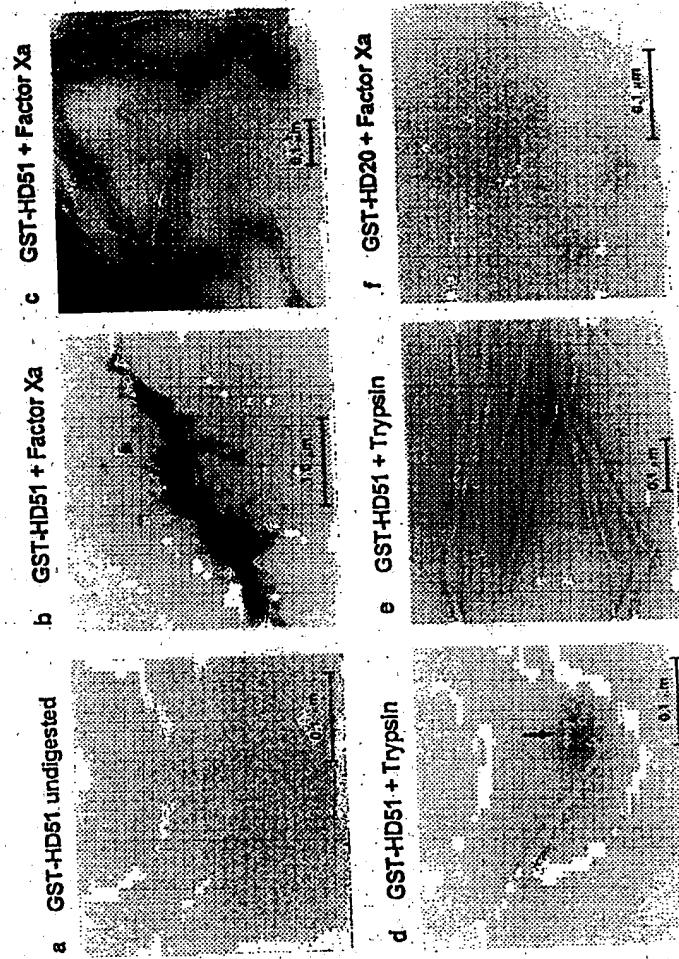


Figure 4  
4/12

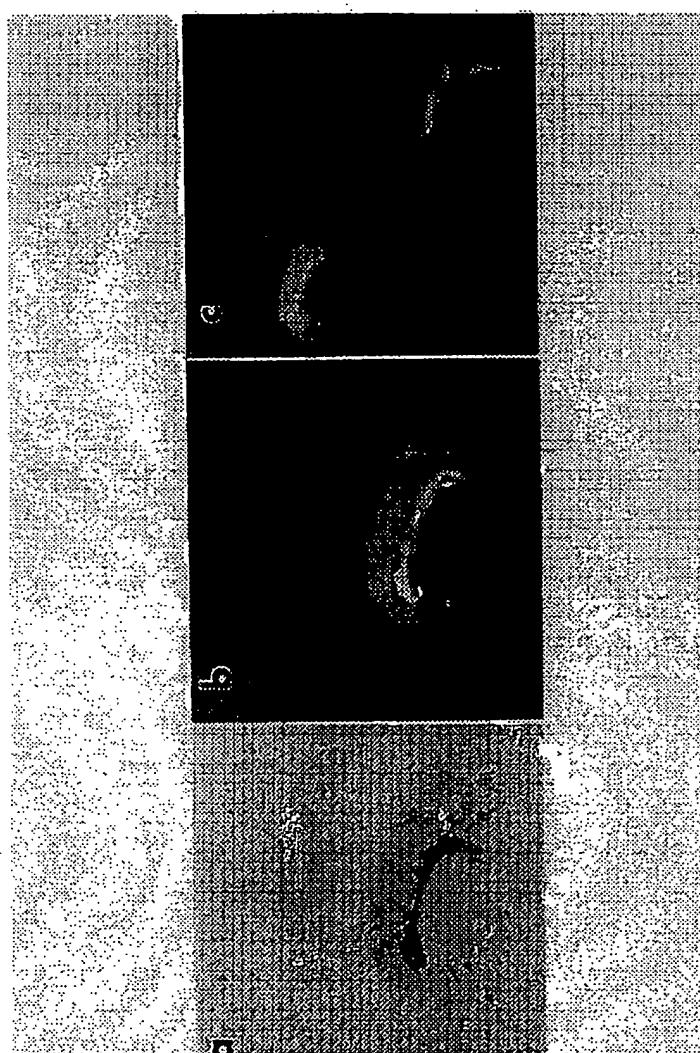


Figure 5  
5/12

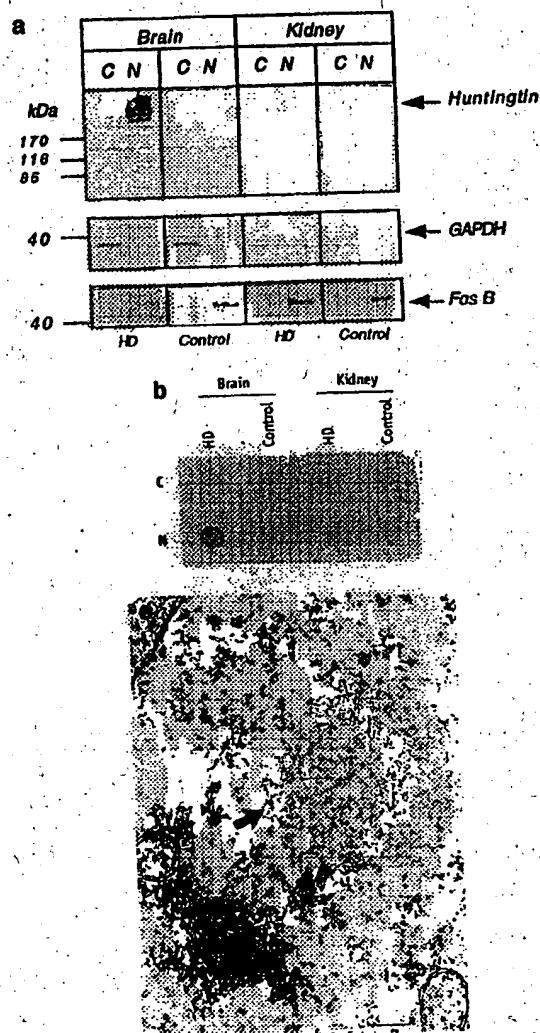


Figure 6  
6/12

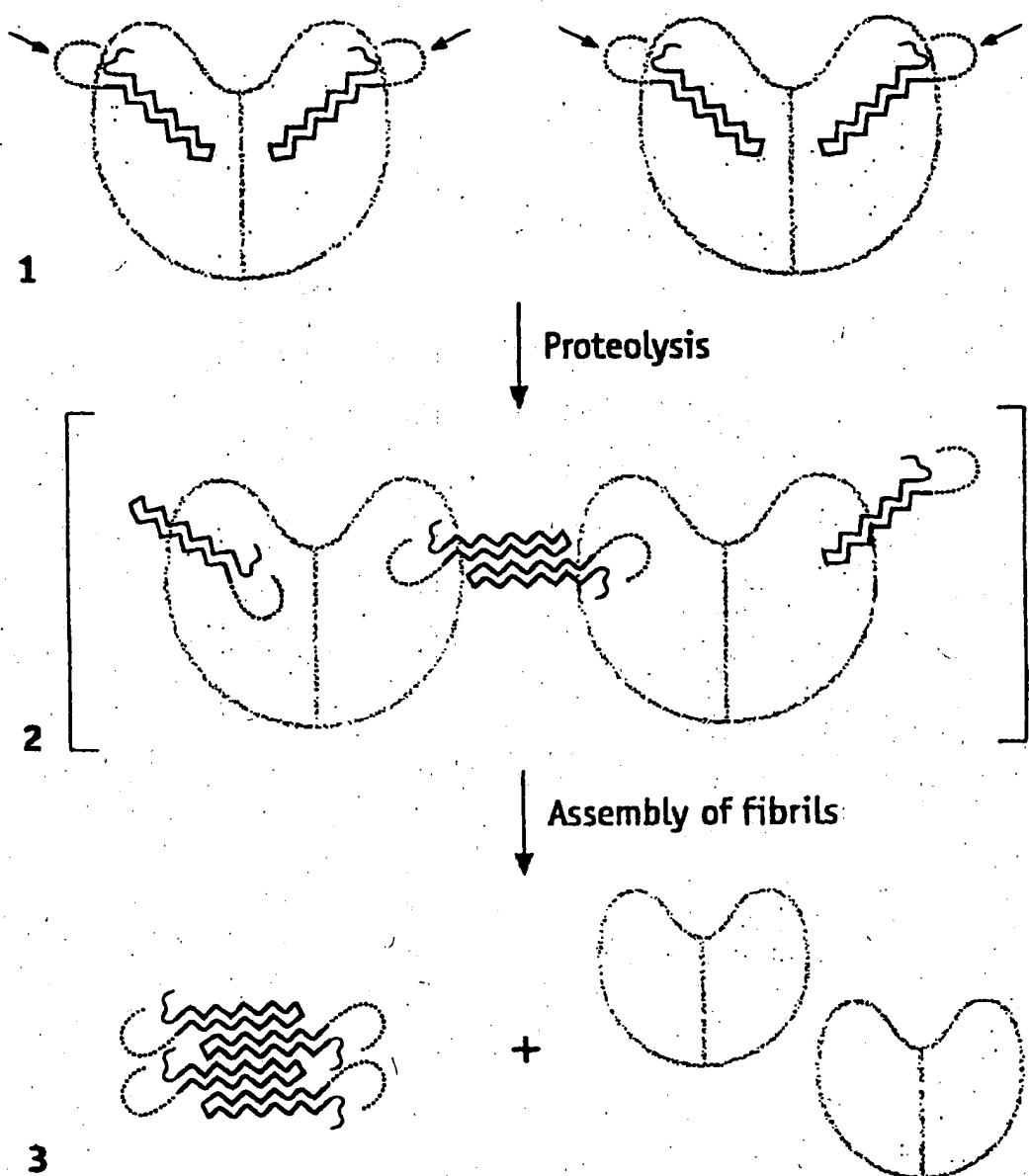
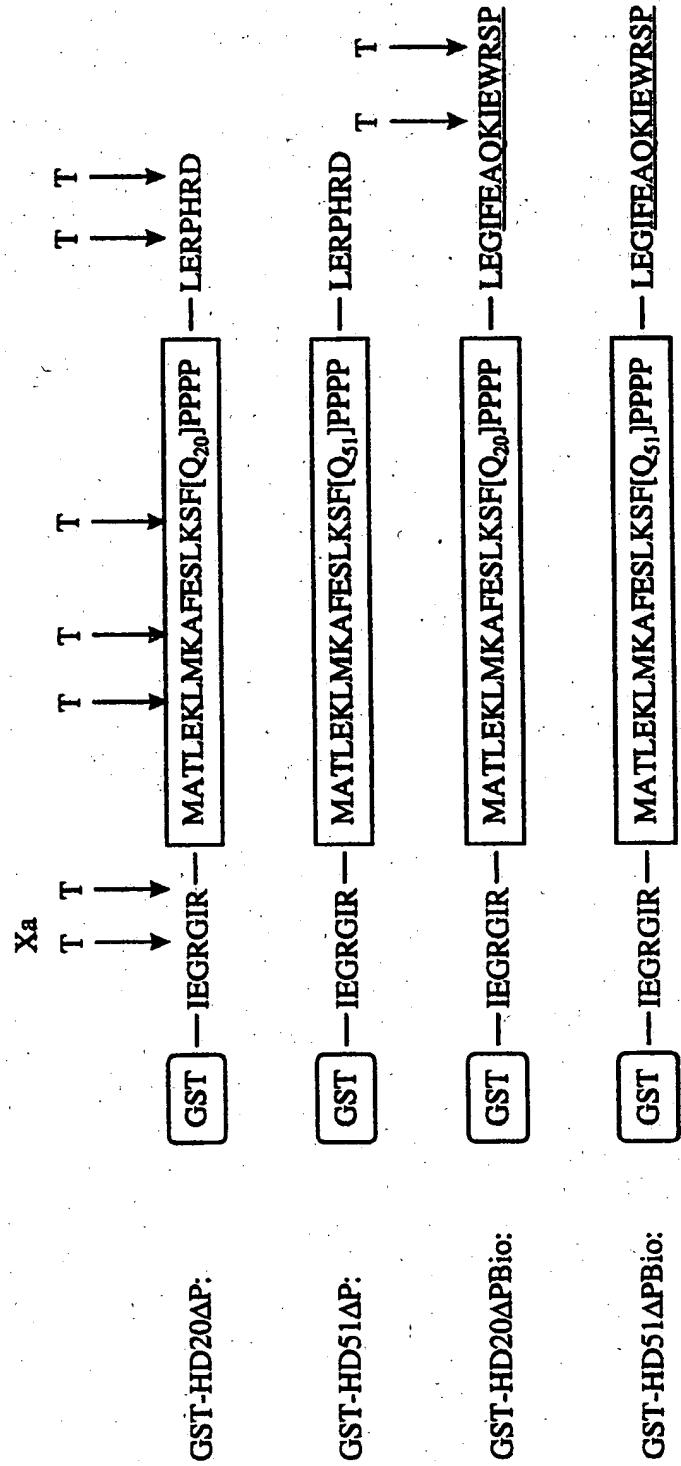


Fig. 7  
7/12



8

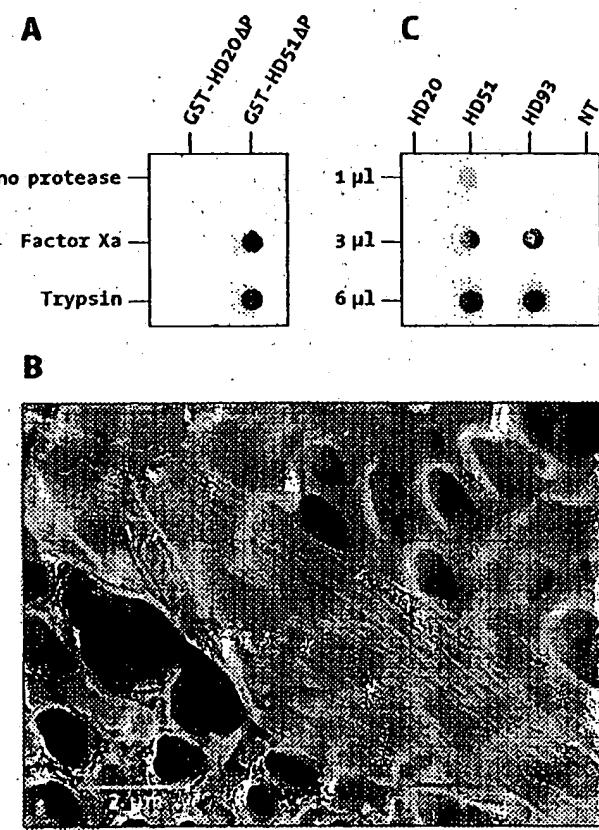


Figure 9

9/12

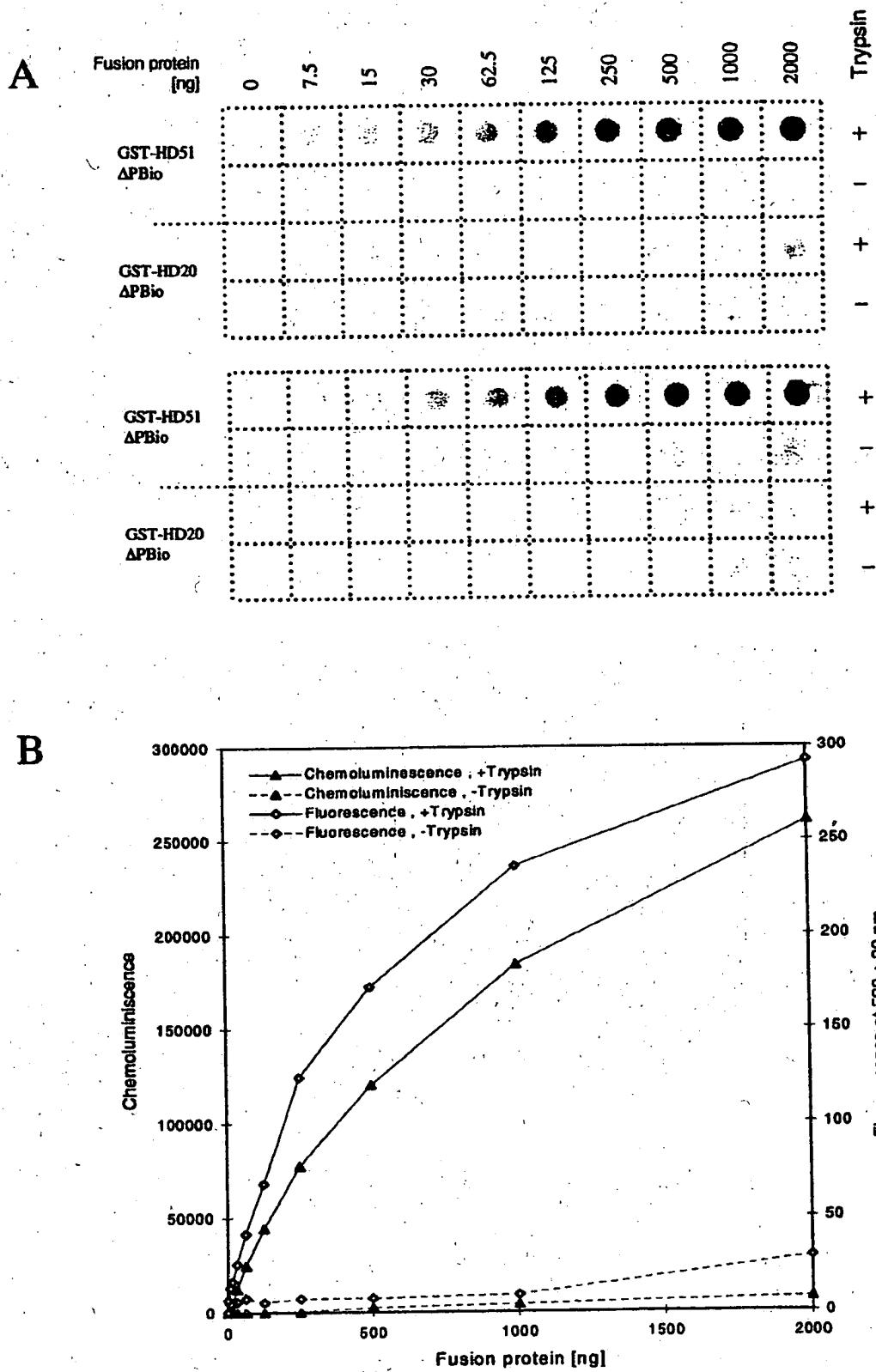
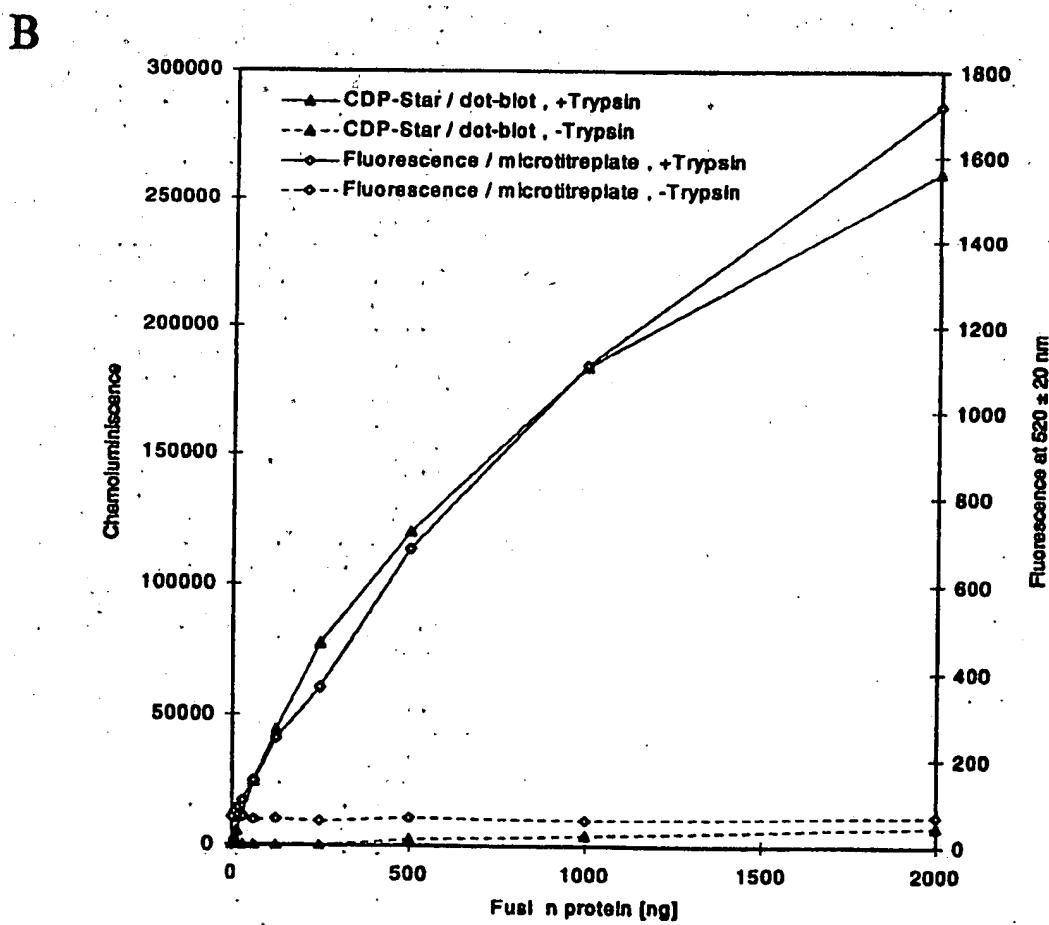
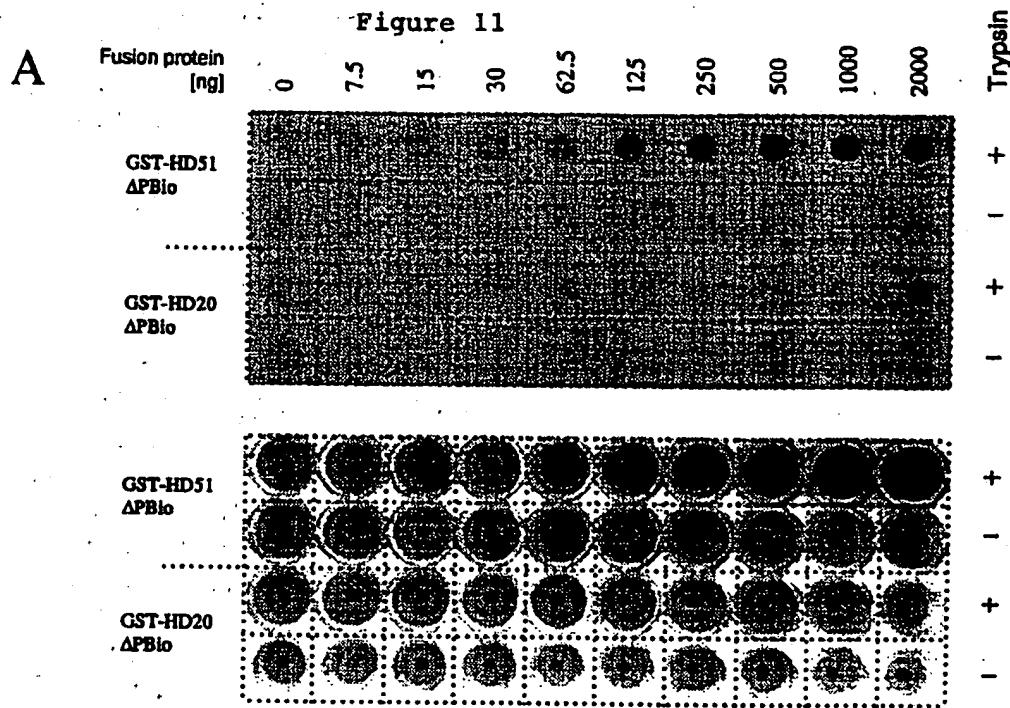


Figure 10  
10/12

Figure 11



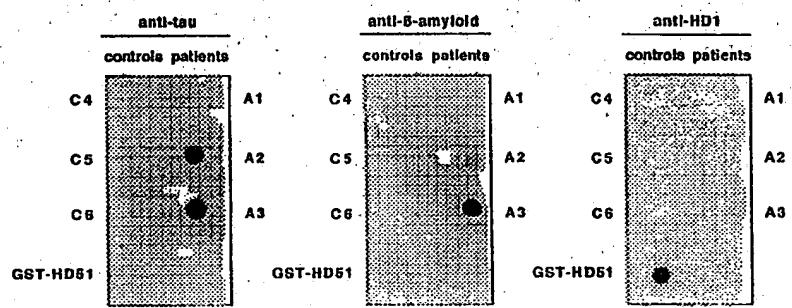


Figure 12

12/12